1872 PRIFYSGOL ABERYSTWYTH UNIVERSITY

IBERS
Athrofa y Gwyddorau Biolegol, Amgylcheddol a Gwledig
Institute of Biological, Environmental and Rural Sciences

Miscanthus biomass quality for conversion: exploring the effect of genetic background and nutrient limitation on the cell wall

Rosario Iacono

This dissertation is submitted for the degree of Doctor of Philosophy

January 2021

Word Count of thesis: 64,726

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Candidate name

Signature:

Date 07/01/2021

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Where *correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).

Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signature:

Date 07/01/2021

[*this refers to the extent to which the text has been corrected by others]

We must trust to nothing but facts: These are presented to us by Nature, and cannot deceive. We ought, in every instance, to submit our reasoning to the test of experiment, and never to search for truth but by the natural road of experiment and observation.

Antoine Lavoisier, Preface to Elements of Chemistry, 1789

Acknowledgements

Here I would like to acknowledge the people who made the completion of this thesis possible.

First of all, I would like to express my sincere gratitude to my supervisors. To Dr Maurice Bosch, for his unconditional trust and unshakable patience. He promoted my personal and professional development. To Dr Gordon Allison for his help and advice and the inspiring conversations. I would also like to thank the people that shared with me the daily work at IBERS. In particular, Agnieszka Gladala-Kostarz, Marina Muñoz Triviño, Ludi Wang, José Carli, Mrs Emma Timms-Taravella, Dr Samantha Gill, Mr Simon Betts and Ms Sue Dalton for creating a supportive and productive work environment. I would like to say thank you to Dr Ricardo da Costa, Dr David Walker and Dr Rakesh Bhatia for their help and advice on the analytical methods. Thanks to Alvaro Garcia for the friendship and the contribution to the harvest of the samples in chapter six.

I am grateful to the student community at Aberystwyth University for making the last four years unforgettable. In particular, Laura Cammarisano, a fantastic scientist and a great friend, for the patience, the advice and the time we shared. Thanks to Antonella Iurato, Gaetano Verdoliva and Gianmarco Sanfratello for the time together, the support and encouragement.

On a more personal level, I want to thank Irene Odoardi who supported my decision of return to academia, encouraged and helped me for the first two years of this journey and to Luciano Canova, for being a great friend who reminded me who I am and inspired my decision of starting a PhD.

A special thanks to Nele, an incredible person and a fantastic partner. For all the moments we shared, the help, the inspiring conversations, for her critical views, but most of all, for never let me forget who I am. Last but not least, thanks to my mum Carmela and my sister Maria who accepted my decisions and supported my commitment to science. A thought to Nunzio, my dad and Maria, my grandmother, whose memory and advise was with me all along this journey.

Finally, I would like to thank the BBSRC iCase and Drax for the financial support.

Summary

Miscanthus is a promising dedicated biomass crop for the bio-industry. The genus includes 16 species and recently interspecific hybrids of *Miscanthus* with increased biomass yield have been created. However, our ability to use *Miscanthus* biomass for producing energy and chemicals sustainably and profitably depends on the ability to cultivate it on marginal lands and harnessing its cell wall recalcitrance.

Although cell wall characteristics vary between genotypes, the full range of this variability in the genus has not been fully explored yet. We selected 49 *Miscanthus* genotypes, from 10 different genetic groups, according to the climate and soil conditions of their area of origin to investigate the extent of variability of cell wall characteristics across the *Miscanthus* genus.

We know that, besides genotypic variation, cell wall properties are also affected by abiotic stresses. However, there is little information about how nutrient limitation, typical in marginal lands, affects the cell wall composition of different *Miscanthus* species and their hybrids. To address this, we selected 12 genotypes from the beforementioned group, with contrasting cell wall structure and quality for conversion, and imposed a nitrogen and phosphorous limitation during their growth for 60 days.

The recalcitrance of cell wall to deconstruction was affected by the organ in a species-specific manner. In general leaf material showed a lower recalcitrance to enzymatic deconstruction. Stem material compared to the leaf material had a higher xylan decoration, a lower ratio between cellulose and hemicellulose and a higher ratio between syringyl and guaiacyl monomers in the lignin. Under nitrogen and phosphorus limitation, there was a species-specific alteration of the saccharification efficiency and the carbohydrate profile of the cell wall.

The outcomes of this work increased our knowledge about cell wall composition in the genus *Miscanthus* and the effect of specific nutrient stress on cell wall composition and recalcitrance. A future for the *Miscanthus* biomass crop can be envisioned where specific nutrient limitation can be used as a treatment to direct the quality of the biomass for conversion.

Contents

Appendixes

List of figures

List of tables

List of Abbreviations and Acronyms

Acronym	Meaning	Page
ABSL	Acetyl Bromide Soluble Lignin	116
ADF	Acid Deterget Fiber	35
ADL	Acid Deterget Lignin	35
AIR	Alcohol Insoluble Residue	27
ANOVA	ANalysis Of VAriance	37
Ara	Arabinose	43
Ara_Enz_p	Enzymatic arabinose release as a percentage of total arabinose	36
AraE	Enzymatic arabinose release as a percentage of total arabinose	35
AXS	Arabino-Xylans	35
BSTFA	N,O-Bis(trimethylsilyl)trifluoroacetamide	106
CADENCOMP	CAnonical polyadic DECOMPosition	
Cel	Cellulose	112
CelHem	Cellulose to hemicellulose ratio	35
Cell	Cellobiose	56
CP		91
CrCel	Crystalline Cellulose	168
CVS	Control Verification Standard	31
CWM	Cell Wall Material	27)
DAPC	Discriminant Analysis of Principal Components	10
DI	De-Ionised	28
DM	Dry Matter	35
DOT	Days Of Treatment	138
DSC	Differential Scanning Calorimetry	70
DSLR	Digital Single-Lens Reflex	102

Acronym	Meaning	Page
DTA	Differential Thermogravic Analysis	72
DTG	Differential ThermoGravimetry	74
DW	Dry Weight	36
DWL	Dry Weight of Leaf	36
DWS	Dry Weight of Stem	36
EG	Emitted Gases	70
EG-FTIR	Emitted Gases - Fourier Transformed Infra-Red spectroscopy	89
EMI/PRI	European Miscanthus	56
Fuc	Fucose	110
Gal	Galactose	43
GC	Gas Chromatography	106
GC-MS	Gas Chromatography - Mass Spectrometry	70
Glc	Glucose	35
Glc_Enz_p	Enzymatic glucose release as a percentage of total arabinose	36
GlcE	Enzymatic glucose release as a percentage of total arabinose	35
GPS	Global Positioning System	8
H:L	Hemicellulose to Cellulose ratio	77
Hem	Hemicellulose	110
HMF	Hydroxymethyl furfural	69
HPAEC-PAD	High Performance Anion Exchange Chromatography - Pulsed Amperometric Detection	30
HSD	Honestly Significant Difference	37
HWSD	Harmonised World Soil Database	8
Hyb.gig	Naturally occurring hybrid	38
hyb.mx	Artificial interspecific Hybrid	38
Hyb.sin	Artificial intraspecific Hybrid	38
K_Lignin	Klason Lignin	35
LA	Leaf Area	136

Acronym	Meaning	Page
LED	Light Emitting Diode	136
LOESS	Local Polynomial Regression	141
Lut	Lutarioriparius	66
M.	Miscanthus	
Man	Mannose	43
ManGlc	Mannose to Glucose ratio in the matrix	48
MLGs	Multi-Linkage Glucans	123
Mns	Monosaccharides	164
MnsT	Total Monosaccharides	163
MW	Molecular Weight	131
N-	Nitrogen limiting conditions	127
n_stems	Number of stems	155
NDF	Neutral Detergent Fiber	35
NIR	Near InfraRed Spectroscopy	33
P-	Phosphorus limiting condition	128
PARAFAC	Parallel Factor Analysis	74
PC	Principal Component	74
PC1	First principal component	147
PC2	Second principal component	147
PCA	Principal Component Analysis	9
PEA	Plant Efficiency Analyser	136
ppm	Parts Per Million	133
PSII	Photosystem II	136
Py-GC–MS	Pyrolisis-Gas Cromatography Mass Spectromery	70
Sacc_Rob	Miscanthus sacchariflorus/robustus	67
sd	Standard deviation	110
sin x sin	M. sinensis intra-specific hybrids	38
Sin_North_Japan	M. sinensis from North Japan	
Sin_South_Japan	M. sinensis from South Japan	

Acronym	Meaning	Page
SPAD	Soil Plant Analysis Development	137
SPAD	Soil Plant Analysis Development	153
SQL 3	Structured Query Language	8
SRS	Sugar Recovery Standard	29
TFA	Tri-Fluoro Acetic acid	104
TG	Thermogravimetry	82
TGA	Thermo Gravimetry	70
TG-DSC	Thermo Gravimetry- Differential scanning Calorimetry	71
TG-FTIR	Thermo Gravimetry-Fourier Transformed Infrared Spectroscopy	71
TG-GC-MS	Thermo Gravimetry- Gas Chromatography - Mass Spectrometry	70
t-test	Student Test	119
wt%	Percentage of weight	68
XG	Xylo-glucan	35
Xyl	Xylose	43
Xyl_Enz_p	Enzymatic xylose release as a percentage of total xylose	36
AraXyl	Ratio between arabinose and xylose in the matrix	48
XylE	Enzymatic xylose release as a percentage of total xylose	35
XylGlc	Ratio between xylose and glucose in the matrix	48

1 *Miscanthus* a dedicated biomass crop: opportunities and challenges

1.1 From the oil-based to the plant-based economy

Modern human civilization is based on the availability of an easily accessible source of energy (Clark, Luque, and Matharu 2012). Since the late 19th century, this source of energy has been fossil carbon. Nowadays, oil is not only the primary source of transportation fuel, but it also represents the primary source for plastics, fibres and colours. There are more than 2,500 products on the market derived from petrol. The chemical industry accounts for about 7–8% of the total consumed liquid and gaseous hydrocarbons. Although essential, oil and other fossil sources are limited. Since the oil crisis in 1970, the notion of the limitedness of fossil reserves generated an increased awareness of the necessity of producing energy and chemicals from renewable sources. Moreover, the prolonged use of fossil fuels, that started in the early 19th century with the use of coal has resulted in anomalies of the natural climatic trends on our planet since the beginning of the 20th century (Stocker et al. 2013). The use of fossil fuels for anthropic activities has resulted in the release of carbon dioxide in the atmosphere that produced an increase in the temperature of the atmosphere, a phenomenon known as global warming. The process was already hypothesised by Svante Arrhenius in 1897 (Arrhenius 1897), who calculated that a reduction of half of the amount of CO2 in the atmosphere would have produced an ice age and a doubling of the amount would have resulted in total warming of 5–6 degrees Celsius. However, in 1897 he thought that the process was negligible, because according to his calculations, at the rate of CO2 emissions in 1897 that would have taken thousands of years. As a consequence of a luxury use of fossil resources, this scenario materialised in less than 100 years with climate warming affecting several

aspects of our ecosystems and human society (Hegerl et al. 2019). Differently from fossil sources, producing energy from renewable sources generates a lower increase in the total amount of CO2 in the atmosphere compared with fossil fuels. Thus, their use could contribute to slow down the effects of global warming. A renewable source of energy, able to completely replace the fossil sources, has to satisfy the needs for the current usage and for future energy demand and for the production of chemicals and materials. In this regard, the attention has focused on the use of plant biomass. Biomass is an important reservoir of organic carbon on Earth that could be used as a source of molecules for the production of fuels and chemicals while reducing the carbon emissions dramatically (Bispo et al. 2017). One of the challenges of our time is to develop a system to effectively exploit biomass as a source of organic carbon. The biomass can be converted to energy and fuels through three main approaches:

The thermal conversion that involves heating the biomass in an oxidative environment to release the energy in its chemical bonds. It is one of the oldest conversion methods known by humanity. In this case, the main product is heat that can be converted to electric power. Because of the dispersion of energy and the low flexibility of the process, it has low efficiency. This conversion route alone does not obviously allow the use of biomass to replace fossil sources since there is no report of the possibility of using it for producing chemicals.

Thermochemical conversion involves the conversion of biomass into a variety of liquid and gaseous fuels upon rapid heating to very high temperature under anoxic atmosphere. The process has been known since 3000 BC (Fahmy et al. 2020). More recently, it has been improved with the developments of new technologies such as microwave pyrolysis (Zhang et al. 2017), solar pyrolysis (Bashir et al. 2017; Li et al. 2016), plasma pyrolysis (Tang and Huang 2005; Huang et al. 2016). It allows for the production of biodiesel and

biogas that could partly replace transportation fuels and be used as a source of molecules for the chemical industry. Recently it has been proposed for the direct production of hydrogen from biomass (Tanksale, Beltramini, and Lu 2010; Waheed and Williams 2013).
Biochemical conversion involves the conversion of biomass into chemicals and fuels upon enzymatic degradation and biochemical digestion with various organic agents (e.g. yeast). The main product of this process is ethanol. It has already been implemented for transportation fuels, although it is still used in a blend with oil fuel (Demirbas 2005).
The chemical conversion that uses chemical agents, either acids or bases to deconstruct the biomass and release molecules that can be used for the production of high-value chemicals and fuels (Jiang et al. 2016)
To implement these conversion methods, the concept of bio-refinery, as opposed to the oil refinery, has arisen over the last few decades. The concept of biorefinery describes the processing steps that will use biomass as feedstock to produce a wide range of chemicals, fuels and bio-based materials, that can be used as industrial intermediates or sold directly to consumers. Biorefineries are believed to have the potential to become the world's primary method for fuel and chemical generation. The concept of biorefinery was developed in the early 1990s and since then has evolved. Up to this moment, three main types of bio-refineries have been developed (Clark and Deswarte 2008). Namely:

- Phase I biorefineries that utilize only one feedstock material, have fixed processing capability and produce a single primary product. A typical example of this type of biorefinery are the numerous biorefineries in Europe that use vegetable oil, mostly rapeseed oil, to produce fixed amounts of biodiesel and glycerin through transesterification. Their lack of flexibility makes it almost impossible to recover investment and operation costs (Wellisch et al. 2010; Clark and Deswarte 2008).

- Phase II biorefineries, similar to the Phase I biorefineries, that use only one feedstock, although they can produce various products. Two examples of this type of biorefinery are the Novamont plant in Italy and the Roquette site of Lestrem in France. The former uses corn starch as a feedstock to produce a range of chemical products including biodegradable polyesters (Origi-Bi) and starch-derived thermoplastics (Mater-Bi). The latter produces a multitude of products including polyols, native and modified starches, proteins and derivatives, cyclodextrins, organic acids and resins.
- Phase III biorefineries are biorefineries that can utilize various types of feedstocks and processing technologies, and produce multiple types of products. They are the most advanced types of biorefineries and are also referred as zero-waste biorefineries (Clark and Deswarte 2008). An example of this type of biorefinery is the Processum Biorefinery Initiative in Sweden. It is a cluster of companies that work together to convert recycled materials, waste, wood and agricultural feedstock to a vast range of chemicals and fuels.

According to the feedstock and conversion process used, there are five Phase III Biorefinery systems (Clark and Deswarte 2008; Takkellapati, Li, and Gonzalez 2018):

1. Lignocellulosic biorefineries that use an array of processes to convert lignocellulosic biomass such as wood, straw and corn stover such as wood, straw and corn stover into a range of products after fractionation. For instance, ZeaChem, a company based in Port of Morrow (Boardman, Oregon, US), uses a hybrid biochemical and thermochemical process to obtain bio-based fuels, C2 chemicals (acetic acid, ethyl acetate, ethanol and ethylene) and C3 chemicals (propionic acid, propanol and propylene) from locally sourced woody biomass and agricultural residues.

2. Whole-crop biorefineries where the feedstock is represented by a whole crop, usually a cereal grain crop. The plant is first separated into corn and straw, and the corn subsequently converted into starch. In this type of refinery, starch is hydrolysed into glucose and constitutes the base for the production of chemicals. The straw is processed in a lignocellulosic biorefinery.
3. Green biorefineries where the feedstock is fresh plant biomass such as grass, green crops or other plant species.
4. Recently, the National Renewable Energy Laboratory (NREL), introduced the concept of a two-platform biorefinery. The two-platform biorefinery includes a sugar platform that employs biochemical conversion methodologies (fermentation of sugars obtained from biomass), and a syngas platform that uses thermochemical methods (gasification of biomass to produce syngas) (Davis et al. 2018).
5. Marine biorefineries that are characterised by using micro- and macroalgae (seaweed) as feedstock. They are younger and less developed than the other biorefineries based on terrestrial feedstock. An example of this type of biorefinery is PlantSea, a startup based in Aberystwyth (Wales, UK) that produces bioplastic from the seaweed from the Welsh coast.

The profitability of a biorefinery is based on maximising the production of high-value low volume and low-value high-volume products while minimizing the waste by converting low-value high-volume intermediates into energy (Cheng and Wang 2013). The first step towards the production of high-value products in a biorefinery is the conversion into platform chemicals. Platform chemicals are chemicals that can serve as a substrate for the production of various other higher value-added products. Ethanol, furfural, hydroxymethylfurfural, 2,5-furan dicarboxylic acid, glycerol, isoprene, succinic acid, 3-hydroxypropionic acid/ aldehyde, levulinic acid, lactic acid, sorbitol, and xylitol have been

identified as the 12 potential platform chemicals that can be obtained from biomass. Figure 1.1 shows the chemical structures of these blocks. They were chosen based on chemicals used in traditional petrochemical industry flow-charts (Werpy and Petersen 2004). The initial screening started with 300 candidate molecules and was reduced to 50 using three criteria:

Direct replacement. According to this criterion molecules, were chosen that directly competed against existing products and chemicals derived from petroleum. An example of a building block selected with this criterion are propylene or lactic acid, used for the production of acrylic acid.

Novel properties. Molecules that possess new and improved properties for replacement of existing functionality or new applications were chosen according to this criterion. For instance, lactic acid, used for the production of polylactate was selected according to this criterion.

Potential utility. Some chemicals selected were the basis of a diverse portfolio of products from a single intermediate, for example: succinic, levulinic and glutamic acids, glycerol.

Subsequently the group of molecules selected were reviewed for chemical functionality and potential use and 30 molecules were selected for a second round of screening that led to the above mentioned 12 sugar derived building blocks.

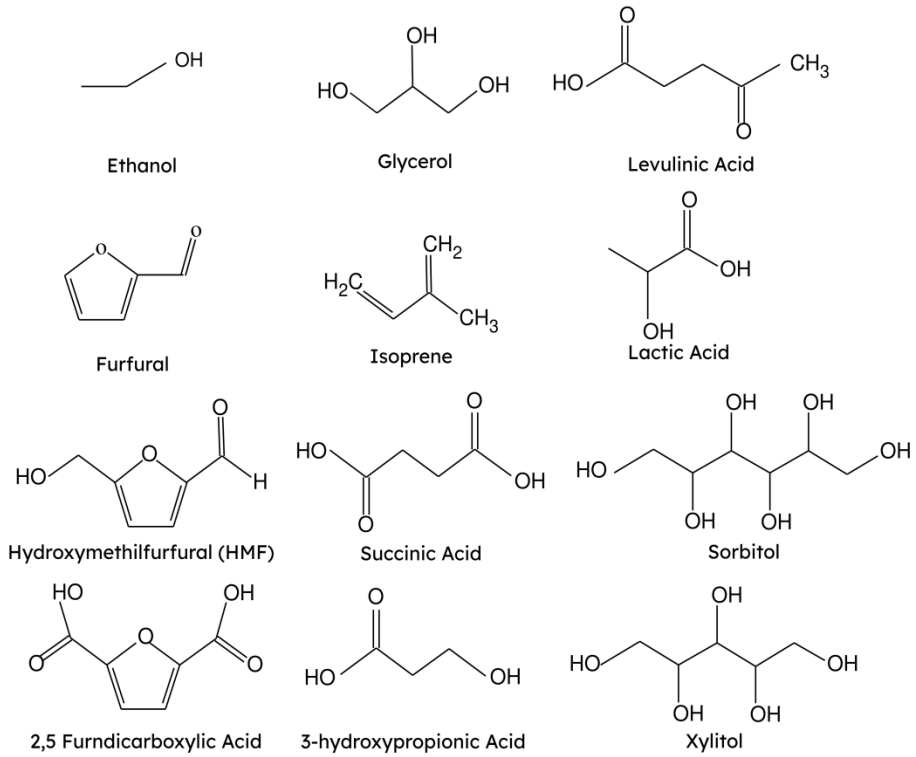


Figure 1-1 Simplified bond-line formulas of the 12 potential platform chemicals that can be produced from lignocellulosic biomass.
Structures were created using Chemtool 1.6.13

1.2 Dedicated crops for biomass production

The first generation of biorefineries was designed to use feedstocks such as corn starch or sugar cane that are renewable, but that also have feed/food uses. The increase in the level of production of these biofuels along with the rise of human population raised concerns about the sustainability of this type of biorefineries. This concern resulted in the development of the second and third generation biorefineries. Different species of plants have been considered for dedicated production of biomass. The choice of the best species for biomass production cannot be univocal and depends on the specific geo-climatic, social characteristics of each area and the conversion process used by the biorefineries in the area (Quinn et al. 2015). In Northern European environments, one of the best candidates to become a dedicated biomass crop are species of perennial grasses belonging

to the genus *Miscanthus* (Lewandowski et al. 2000). They have the potential of generating high-yield thanks to the ability to maintain their C4 metabolism in temperate environments and they require minimal soil tilling, fertilisers and pest control (Varnero, Urrutia, and Ibaceta 2018). When grown on marginal lands, they can improve the structure of the soil, increase water-holding capacity and reduce run-off and erosion (McCalmont et al. 2015). *Miscanthus* is a genus of tropical monocotyledons belonging to the *Poales*. The genus includes 16 confirmed species distributed on a broad range of latitudes and altitudes in east Asia (Clayton et al. 2006; Slavov et al. 2013; "The Plant List - Version 1" 2010). The list of accepted species in the genus *Miscanthus* can be found in Table A 1. In areas where *Miscanthus* sinensis and *Miscanthus* sacchariflorus cohabit, sterile interspecific hybrids between these two species have been reported (Tamura et al. 2016; Uwatoko et al. 2016), named *Miscanthus* x *giganteus* (Hodkinson and Renvoize 2001). Until recently, the use of *Miscanthus* spp. as a biomass crop was related to clones of that original hybrid, obtained by rhizome splitting. The use of clonal propagation has represented a bottleneck in the development of this species as a biomass crop because of the difficulty to scale up the production of seedlings and the limited genetic variability available for breeding. Consequently, the attention has now shifted to the use of natural accessions for the creation of new interspecific hybrids similar to the ones reported in natural environments (Kalinina et al. 2017; Clifton-Brown et al. 2019).

1.3 Biomass quality and cell walls

1.3.1 Cell wall composition and structure

Plant biomass material is constituted by cell solubles and cell wall. The cell wall represents between 45 and 65% of the biomass in dry weight (Robbins and Moen 1975) of a plant and it is an important source of molecules for the production of the renewable molecules described in section 1.1.

The cell wall is a complex composite structure that surrounds all the cells of plants (Albersheim et al. 2011). Its primary functions are the maintenance of structural integrity by resisting internal hydrostatic pressures, the determination of the shape of the cell and the support for the whole plant (Tonzig and Marré 1983; Houston et al. 2016; Albersheim et al. 2011; Alberts et al. 2002). In addition to this function, the cell wall is flexible during cell division, enables differentiation by providing the biochemical scaffold, and acts as pathological and environmental barrier that defends against stress (Houston et al. 2016). A new cell wall is generated at the interface of two separating cells. Thus, every facet of the polyhedral wall is laid down at a different time and reflects the events that happened at different times of cell development (Albersheim et al. 2011). As the development of the cell proceeds, the cell starts its expansion, and concomitantly new cell wall material is juxtaposed internally to the existing cell wall and expanded accordingly. Once the cell completes its extension, it starts to specialise. The specialization of a cell to a particular function usually involves modifications of the cell wall to allow the cell to play its specific function. The mechanisms that regulates this transition, which is at the origin of plant developmental flexibility to environmental conditions, is still a challenging field for research in developmental biology (Pierre-Jerome, Drapek, and Benfey 2018). The process is controlled by positional cues generated by signaling and regulatory feedback (Berg et al. 1997). For instance, stomatal guard cells have asymmetrically thickened cell walls to allow them to reversibly change shape in the process of stomatal closure and opening. Cells can specialise to three main types of plant tissues, the dermal tissue system, the ground tissue system and the vascular tissue system. Each organ of the plant is made by a combination of these three main tissue types. Consequently, the characteristics of cell wall material from each organ depend on the composition in terms of tissue types (Tonzig and Marré 1983; Alberts et al. 2002).

Although new cell wall material is apposed continuously internally on the cell wall, it is the convention to identify three layers in the cell wall. A middle lamella formed just after cell division, a primary cell wall able to allow the expansion of the cell and a secondary cell wall more rigid and typical of cells specialised to particular functions. The primary cell wall is composed of three classes of structural polysaccharides, cellulose, hemicellulose and pectins, and sometimes they also contain structural proteins. The secondary cell wall is also containing lignin, a complex aromatic polymer. The relative abundance and the interlinkage between the components of the cell wall vary between organs, tissues and plant phyla (Petit et al. 2019; Fangel et al. 2012).

The main distinction is made between the structural model that describes the cell walls in dicotyledons and most of the non Poales orders of monocotyledons (including the orders *Asparagales*, *Liliales*, *Dioscoreales*, *Pandanales*, *Petrosaviales*, *Alismatales*, *Acorales*, *Arecales*, *Commelinales* and *Zingiberales*) characterised by a Type I cell wall and the order *Poales* of the commelinid monocotyledons (including *Poaceae*, *Cyperaceae*, *Bromeliaceae*), characterised by a Type II cell wall (The Angiosperm Phylogeny Group 2016; Carpita and Gibeaut 1993). These groupings are arbitrary, and a more continuous variation between species can be imagined. For the aims of this thesis, the description of the cell wall composition and structure model will refer to the cell wall of the *Poales* that includes the genus *Miscanthus*. The model that best describes the arrangements of the polymers in the cell wall of *Poales* is similar to the type II cell wall model, as described by (Carpita and Gibeaut 1993), with some differences that will be described below. The type II cell wall structure is composed of cellulose microfibrils embedded in a matrix made of glucuronoarabinoxylans (GAXs), linear chains of (1+4) P-*O*-xylose with single arabinose unit at the C0-3 positions and, less frequently, single glucosyluronic acid (GlcA) units at the C0-2 of the xylosyl units (Carpita and Gibeaut 1993). A diagramatic representation of

the type II cell wall as described can be found in figure 1.2. In the figure the glycan units are represented according to the symbol nomenclature for graphical representations of glycans (Varki et al. 2015; Neelamegham et al. 2019). In the following the peculiar structure of each component and their interrelation will be discussed.

Cellulose is the most stable cell wall polysaccharide and the major structural or load-bearing component. In the *Poales* its structure is not different from the one in other plants. Cellulose is a simple continuous unbranched polysaccharide of beta-1,4-linked glucopyranosyl residues (Fig. 1.2 - A). The beta-1,4-linkage requires that alternated glucosyl residues in the molecule are rotated by 180° to each other. Therefore, the molecule of cellulose can be imagined as formed by repeating units of cellobiose. Cellulose molecules are characterised by the ability to form intramolecular and intermolecular hydrogen bonds. In the cell wall, 30 to 50 cellulose chains are held together by hydrogen bonds to form cellulose microfibrils (Fig. 1.2 - B). Microfibrils are 5-15 nm wide and are spaced 20-40 nm apart (McCann, Wells, and Roberts 1990). Cellulose fibrils organised in this way form a resistant crystalline structure that confers to the cell wall tensile strength and stiffness (Gomez, Steele-King, and McQueen-Mason 2008). In the fibrils, cellulose can be found in its crystalline as well as amorphous status. It has been observed that the two statuses are distributed along the fibrils in a periodic disorder (Nishiyama et al. 2003).

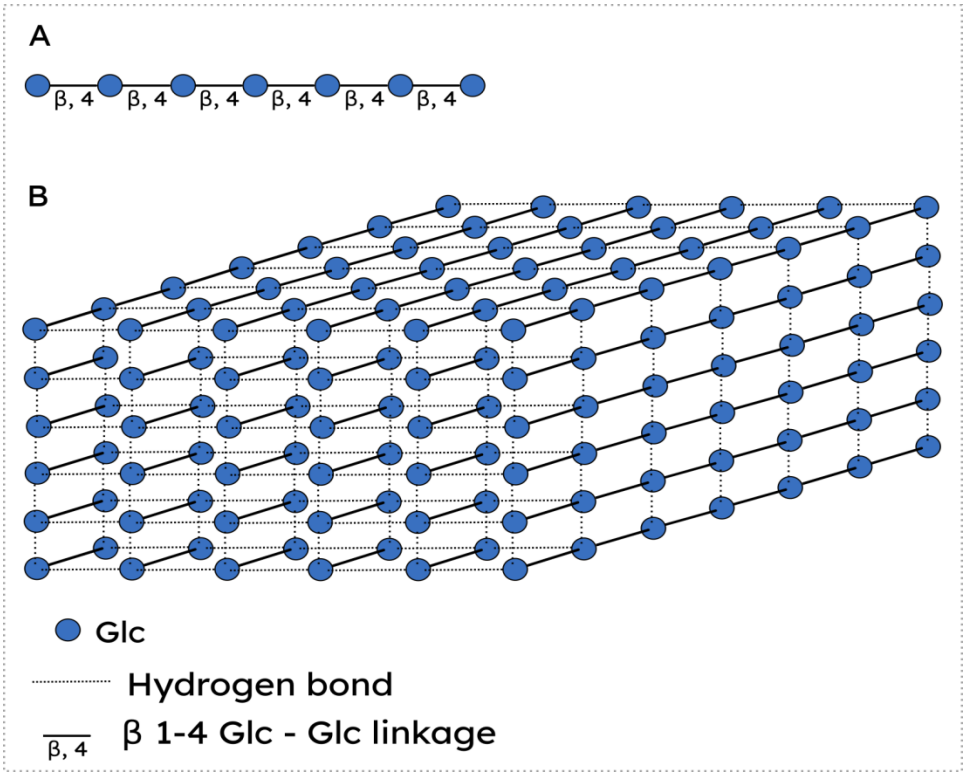


Figure 1-2 Diagrammatic representation of cellulose.
Glycan units are represented according to the symbol nomenclature for graphical representations of glycans, A: Cellulose molecule. A molecule of cellulose is a linear sequence of glucose residues (Glc) connected through 1,4-beta linkages. B: Arrangement of cellulose molecules in a crystalline portion of a cellulose microfibril. The long molecules of cellulose are interlinked by hydrogen bonds (dashed lines)

The interlinking between microfibrils is performed by hemicellulose and pectins. In the *Poales*, the interlinking polymers are glucuronoarabinoxylans (GAXs) and (1,3) (1,4)-beta-D-glucan (mixed-linkage glucan or MLG) (Buckeridge et al. 2004). GAXs are linear chains of (1-4) beta-D-xylose with single arabinose units at the C-3 positions and, less frequently, single glucosyluronic acid (GlcA) units at the C-2 positions of the xylosyl units (Carpita and Gibeaut 1993). In *Poales*, in the side chains arabinofuranose (Araf) residues can be substituted with an Araf dimer or a xylanfuranose-arabinose dimer (Peña et al. 2016). The reducing end of xylans in *Poales* contains branched and unbranched xylosyl residues substituted with Araf (Fig. 1.3 - a), Xylf (Fig. 1.3 - b) and 4-*O*-Methyl glucuronic acid (MeGlcpA) (Fig. 1.3 - c) (Ratnayake et al. 2014; Peña et al. 2016). Arabinosyl

substitution of grass xylan can vary from molar ratios of 1:2 Ara: Xyl to levels of 1:20 or 1:30 depending on the developmental stage of the specific commelinid species and tissue considered. Substitution patterns on the xylan backbone dictate how strongly they can hydrogen bond to other wall polysaccharides, mainly cellulose and to other xylans, influencing structural properties of the wall (Grantham et al., 2017). Another unique characteristic of xylans in monocotyledons is the substitution of alpha-(1,3)-Araf residues with ferulic acid (FA) and, to a lesser extent, p-coumaric acid (pCA) (Ishii and Hiroi, 1990). FA is bound to xylans through an ester bond. FA is also associated with lignin in the secondary cell wall, and the formation of FA dimers links hemicellulose and lignin in the secondary cell wall. FA dimers in the cell wall of *Poales* were discovered as the 5-5'-dehydrodimer, isolated after alkaline hydrolysis of the cell wall polysaccharides (Harris and Hartley, 1976; de O. Buanafina, 2009). Since then, several dehydrodimers, trimers and tetramers have been identified and characterized in different species of poales. The role of *p*-CA decoration of xylans is not clear. It is believed that it plays a catalytic role in lignin biosynthesis. *p*CA acts as a radical transfer system, although it does not become part of the radical-mediated cross-coupling reactions that form the growing lignin polymer (Hatfield, Marita, and Frost 2008).

MLGs are unbranched and unsubstituted chains of beta-glucopyranosyl monomers linked through (1,3) and (1,4) linkages. MLGs are particularly abundant in the primary cell wall and consequently in young tissues. They are part of the interlinking structure, and they are linked to cellulose (Kiemle et al. 2014). In contrast with the fact that MLGs are typical of a recently evolved group of plants, they can also be found in lower vascular plants, such as the horsetails, in algae and in fungi.

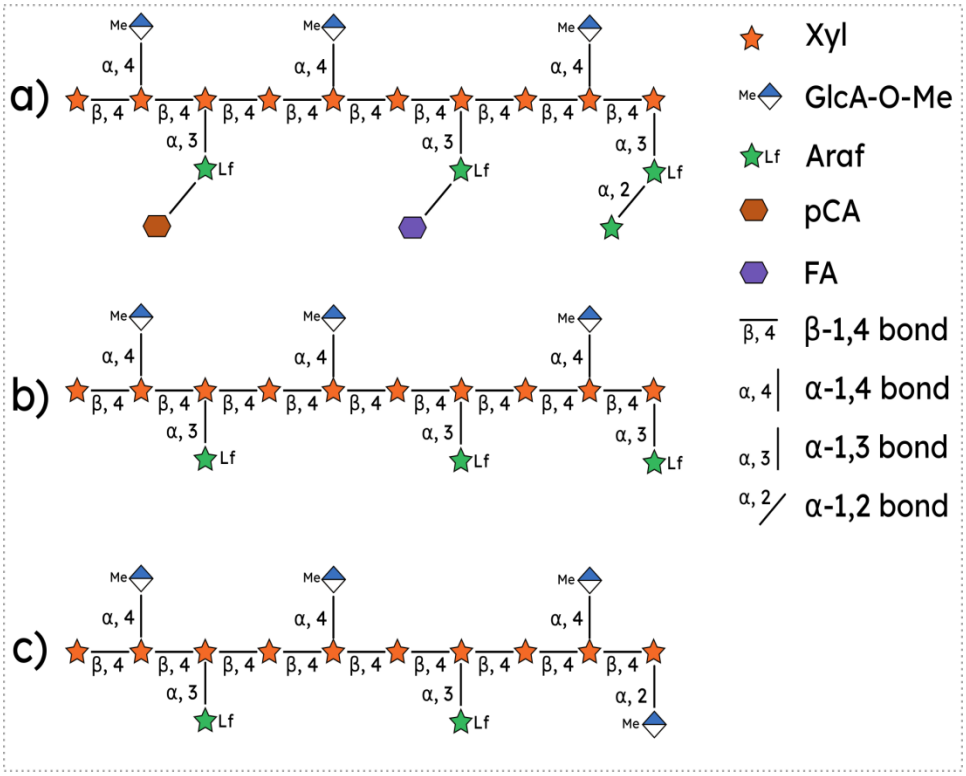


Figure 1-3 Diagramatic representation of glucuronoarabinoxylan (GAXs) structure in Type II cell walls of Poaceae.
Glycan units are represented according to the symbol nomenclature for graphical representations of glycans. GAXs are linear chains of (1-4)-D-xylose (Xyl, yellow star) with single arabinose units (Araf, green star) at the C-3 positions and, less frequently, single glucosyluronic acid (GlcA-O-Me, blue and white divided diamond) units at the C-2 positions of the xylosyl units. The reducing end contains branched and unbranched xylosyl residues substituted with Araf (a), Xylf (b) and 4-O-Methyl glucuronic acid (c)

Cell wall porosity is a physical property of the cell wall defined by measures of pore size and distribution in the structure which is regulated through pectins. Pectins are present in the monocotyledon cell walls in the proportion of 2–10% (Voragen et al. 2009). In addition to porosity, pectin influences various cell wall properties such as surface charge, pH, and ion balance and therefore is of importance to the ion transport in the cell wall. Pectins have been identified as having a crucial role in maintaining cell wall integrity under stress (Hahn, Darvill, and Albersheim 1981; Nothnagel et al. 1983; Wu, Bulgakov, and Jinn 2018).

The main types of pectin in monocotyledon cell walls are polygalacturonic acid (PGA), which is a homopolymer of (1,4)-alpha-D-galactosyluronic acid (GalA), rhamnogalacturonan I (RG I)(Fig. 1.4 - a), which is a heteropolymer of repeating (1,2)-alpha-L-rhamnosyl (Rha)-(1,4)-alpha-D-GalA disaccharide units (Fig. 1.4 - b) and rhamnogalacturonan II (RGII) which is a heteropolymer of repeating (1,2)-alpha-L-rhamnosyl (Rha)-(1,4)-alpha-D-GalA disaccharide units usually decorated on their C-4 positions (Fig. 1.4 - c). In contrast with dicotyledons, in monocotyledons, the side chains on the RGII chains lack fucose.

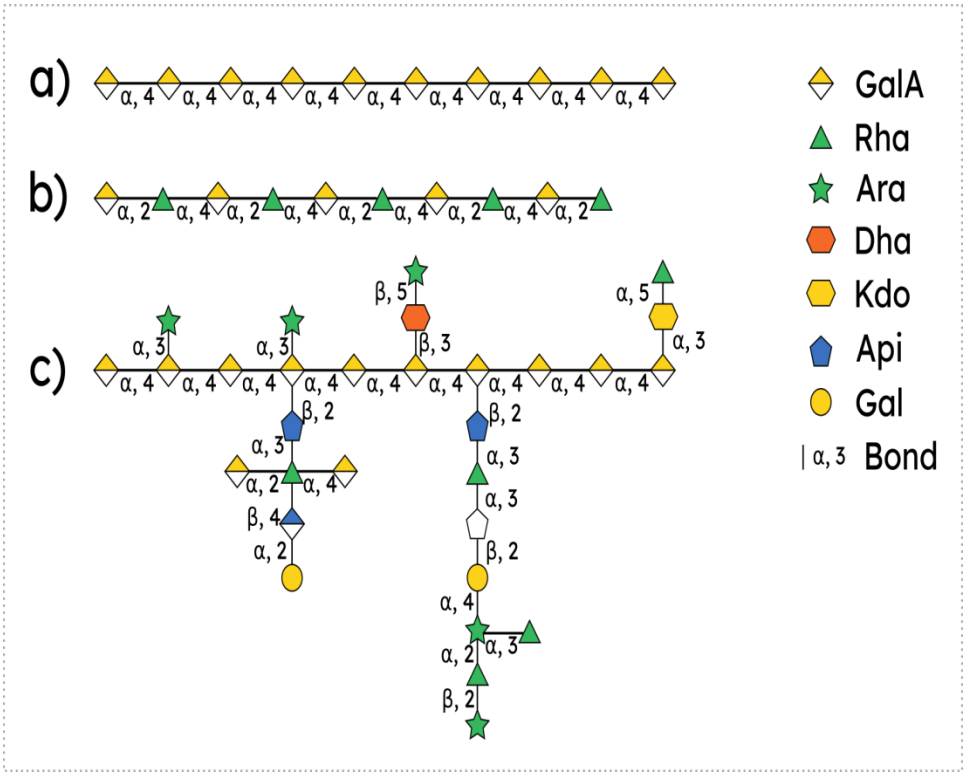


Figure 1-4 Diagramatic representation of pectin structure in Type II cell wall of Poaceae.
Glycan units are represented according to the symbol nomenclature for graphical representations of glycans, labels indicate the covalent bonds between molecules. a): polygalacturonic acid (PGA), which is a homopolymer of (1,4)-alpha-D-galactosyluronic acid (GalA), b): rhamnogalacturonan I (RG I), which is a heteropolymer of repeating (1,2)-alpha-L-rhamnosyl-(1,4) (Rha) and alpha-D-GalA (GalA) disaccharide units, c): rhamnogalacturonan II (RGII) which is a heteropolymer of repeating (1,2)-alpha-L-rhamnosyl (Rha) -(1,4)-alpha-D-GalA disaccharide units usually decorated on their C-4 position

In the secondary cell wall, lignin is deposited within the cell wall matrix. Lignin is a hydrophobic heteropolymeric phenyl propanol material that forces the water out of spaces in the wall matrix as it forms, decreasing the flexibility and permeability (Fig. 1.5). The polymeric structure includes three monomers (monolignols): p-coumaryl, coniferyl, and sinapyl alcohol (Hatfield, Rancour, and Marita 2017; Schäfer et al. 2019). The peculiarity of the type II cell wall resides in the presence of *p*-Coumaric acid linked to the synapyl sub-units [Hatfield, Marita, and Frost (2008);] and the presence of ferulic acid interlinking lignin with the cell wall polysaccharides (de O. Buanafina, 2009). Lignin biosynthesis is controlled by developmental and environmental signals (Campbell and Sederoff 1996; Sederoff et al. 1999). There is evidence that lignin deposition starts around the ferulate decoration of the xylans (Ralph et al. 1994; Grabber, Ralph, and Hatfield 1998). From these nucleation points, it then proceeds to the rest of the cell wall which also explains the ability of grasses to maintain stem flexibility (Hatfield, Rancour, and Marita 2017).

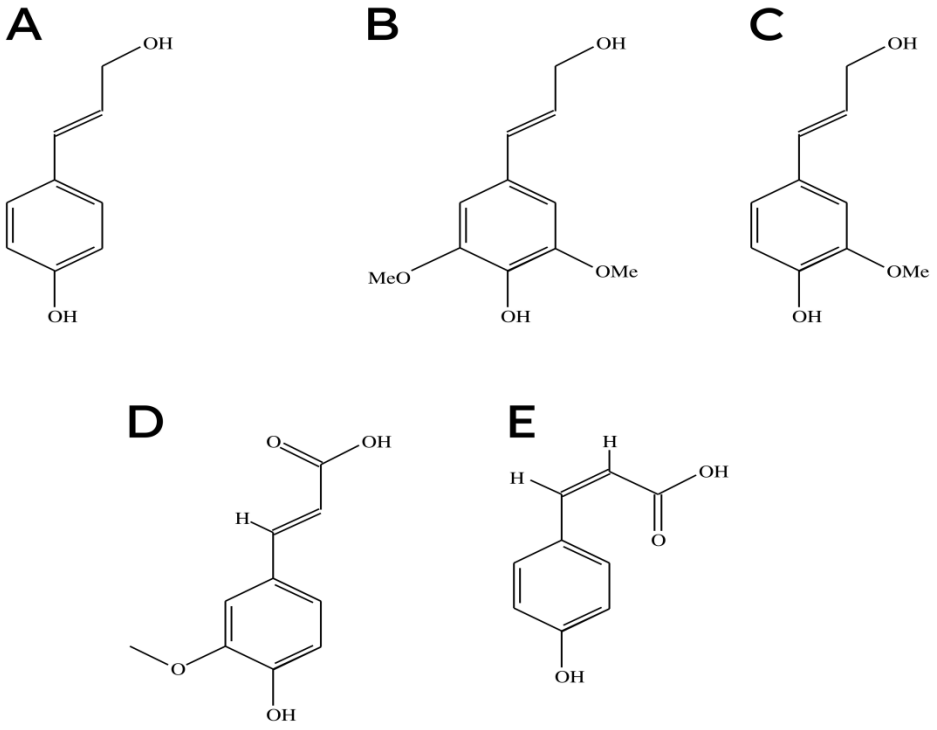


Figure 1-5 Simplified bond-line formulas of lignin monomers (monolignols) in Type II cell wall of Poaceae.
A: *p*-Coumaryl alcohol, B: Synapyl alcohol, C: Coniferyl alcohol, D: Ferulic Acid (FA) and E: *p*-Coumaric acid (*p*CA)

1.3.2 Impact of cell wall composition on conversion

To transform renewable building blocks from plant biomass with the same range and versatility as the petrochemical industry it is necessary to disassemble the biomass and to reassemble the resulting carbon into the platform chemicals described above. The effectiveness of this process is hindered by the cell wall recalcitrance, defined as the resistance of plant cell walls to deconstruction by chemical, enzymatic, and/or microbial routes (DeMartini et al. 2013). Recalcitrance is a complex biomass characteristic resulting from plant ultrastructural, molecular, and chemical features (DeMartini et al. 2013). It has been shown that several compositional features of the cell wall contribute to its recalcitrance (Foston 2014; Himmel et al. 2016). The recalcitrance can be overcome by using pretreatments which include the use of chemicals and high temperatures and pressures to disrupt the cell wall structure prior to enzymatic hydrolysis. A more accurate definition of recalcitrance, that includes this latter aspect of conversion was provided by McCann and Carpita (2015). Starting from the definition of emergence as "the arising of novel and coherent structures, patterns and properties during the process of self-organization in complex systems" provided by Goldstein (1999), McCann and Carpita (2015)defined recalcitrance as "an emergent property of the molecular and structural complexity of plant cell walls *in planta* and how that structure interacts with catalysts during pre-treatments and conversion pathways". In providing this definition they suggested that the recalcitrance is defined by the interaction of the catalyst and their cell wall substrates and thus it is a characteristic that affects the amount of product yield according to the conversion process considered. This in turn affects our ability to define

the of quality of biomass for biorefining, that necessarily has to be addressed in a conversion-specific fashion.

Below the role of the main components of the cell wall on recalcitrance will be described. Cellulose can be a source of glucose upon chemical or enzymatic degradation. However, its level of crystallinity can dramatically reduce the ability to release glucose molecules. Degradation of amorphous cellulose happens at a much higher rate compared with crystalline cellulose during enzymatic digestion, acid hydrolysis, alcoholysis, or acetolysis (Himmel et al. 2016). Comparing cell wall biomass of equal crystallinity, the amount of cellulose is positively correlated with the amount of glucose enzymatically released (Yoo, Kim, and Yoon 2017). On the other hand, the length of the cellulose microfibrils has a negative effect on the amount of glucose released (Lu et al. 2019). Different approaches to increase the efficiency of cell wall degradation have been proposed. Species-specific variation in crystalline cellulose distribution can be used to improve biomass quality of crops. Also, a better understanding of inter-microfibril interactions can contribute to increasing the accessibility of conversion agents into the microfibrils (McCann and Carpita 2015).

Hemicelluloses can be a source of mostly five- but also six – carbon monosaccharides that can be converted into platform chemicals (Takkellapati, Li, and Gonzalez 2018). Hemicellulose increases the recalcitrance to biochemical conversion by acting as a physical barrier and limiting the accessibility of enzymes (Kruyeniski et al. 2019). Its removal, for instance, with an acidic pre-treatment, can increase the ability of enzymatic degradation of cellulose (Santos et al. 2018). The degree of polymerization, substitution pattern and cross-linking initiation sites of hemicellulose are targets for genetic control that could be exploited to reduce biomass recalcitrance (McCann and Carpita 2015).

Lignin can be a source of aromatic compounds for the production of high energy fuels and

platform chemicals (Yang et al. 2019). However, lignin is well known for decreasing the ability to release monosaccharides from cellulose and hemicellulose. This negative role in the conversion is influenced by total lignin content as well as lignin composition/structure, in particular hydroxyl groups content and S and G unit content (Santos et al. 2012; Bhatia et al., 2019). The profitability of biorefineries will rely on our ability to overcome cell wall recalcitrance. This could be achieved either by improving the conversion technologies available as well as by better understanding the biological bases of biomass recalcitrance. Increased knowledge of the processes underpinning biomass recalcitrance could be used to inform breeding programs of dedicated biomass crops. It could also inform the selection of targets for the genetic modification of biomass crops.

1.4 Marginal lands

In a scenario where the world population is increasing, the use of biomass as a feedstock for a biorefinery will only be sustainable in the long run if it does not interfere with food production. Gelfand et al. (2013) demonstrated for the first time that the marginal land in the US could be sufficient to support the biomass production to meet the 25% target of cellulosic biofuel by 2022 (Gelfand et al. 2013). The profit of using marginal lands for bioenergy production is higher when environmental, economic, and benefits are considered (Kang et al. 2013). A similar possibility has been demonstrated in Mediterranean areas in Europe by the results of the project OPTIMA (Optimization of Perennial Grasses for Biomass Production in the Mediterranean Area) (Monti and Cosentino 2015). The definition of marginal land is quite controversial and has been debated since 1932 (Peterson and Galbraith 1932). Today, there is broad agreement in defining marginal lands as areas with marginal agronomic and economic potential for cultivation of food crops and currently not used by agriculture (Peterson and Galbraith

(1932); Dauber et al. (2012); Kang et al. (2013); Carlsson et al. (2017); Mehmood et al. (2017); Pancaldi and Trindade (2020)). Although there is agreement on the definition of marginal land, the factors determining the marginality of lands can be numerous. Lands can be marginalised as a consequence of degradation due to excessive agricultural exploitation, contamination with pollutants, and changes in climate that increase the incidence of drought or flooding. Some characteristics, such as high erodible soil and ecological sensitivity, are determining marginality without directly affecting the possibility of growing crops in the short term. Soil erosion is the factor with the most significant impact in determining marginality (Kang et al. 2013). The marginalisation of soil caused by erosion is often connected with nutrient deficiencies in the soil.

1.4.1 Nutrient stress

The life of plants on earth depends on acquisition of mineral elements identified as plant nutrients from the environment (Sprenger 1839; Reed 1942). Generally, a plant nutrient is defined as a chemical element required for a plant to complete its life cycle, that cannot be substituted fully by any other element, and that is needed by all plants (Barker and Pilbeam 2016). Seventeen elements are considered to meet these criteria: carbon, hydrogen, and oxygen are derived from air or water, and the other 14 are obtained from soil or nutrient solutions that plants acquire predominantly from the soil environment. The six elements required in relatively large amounts are referred to as “macronutrients” (Maathuis 2009). The macronutrients comprise nitrogen (N), potassium (K), calcium (Ca), magnesium (Mg), phosphorous (P), and sulfur (S). The group of elements needed at much smaller concentrations in plant tissues is defined as “micronutrients” or “trace elements”. It comprises chloride (Cl), copper (Cu), manganese (Mn), iron (Fe), zinc (Zn), cobalt (Co), molybdenum (Mo), Boron (B) and Nickel (Ni). Nutrient uptake by plants can be determined by the availability of the nutrient in the environment or by the ability of

the plant to take up nutrients. The nutrient availability in the soil depends primarily on their abundance in a form available to plants. In addition, the concentration of nutrient elements in a form that is available for plants in the soil can fluctuate over time as a consequence of environmental conditions like weather and climate and soil conditions (Maathuis 2013). Moreover, nutrient availability in the soil depends on parent material of the soil (Singh and Schulze 2015), water availability and its movement in the soil layers (Marschner and Rengel, 2011), soil structure (Marschner and Rengel 2011), pH and redox potential especially in the rhizosphere (Neumann and Römheld 2011). White (2011) identified three main characteristics of the plant's ability (both higher and lower plants) to uptake elements in their ionic form: 1. Selectivity. Plants discriminate elements, and the take up is different for each element. Some elements are excluded from uptake. 2. Accumulation. Element concentration in cell sap can rise to a level higher than the one in the external solution. 3. Genotype. Nutrient uptake is under strict genetic control and can be different between species or genotypes. The limitation of one or more nutrient elements, either as a consequence of reduced availability or reduced plant uptake, results in a reduction of metabolic efficiency of plants and a type of abiotic stress referred to as nutrient stress. The symptoms of nutrient stress are specific to the physiological function of each element lacking and its interaction with the other elements in the plant (Maathuis 2009). The most notable effect of nutrients in plant metabolism is on photosynthesis. The nutritional status of the plant affects photosynthesis through several processes including the biosynthesis of the photosynthetic apparatus, element fluxes through membranes, involvement in source/sink balance and other physiological processes (Engels, Kirkby, and White 2011). The parameter that provides an integrated reflection of the nutrient availability in the soil and the ability of the plants for their uptake is the Nutrient use Efficiency (NUE). NUE describes a multigenic complex trait resulting from numerous

interconnected physiological mechanisms modified by several factors (Reich, Aghajanzadeh, and De Kok 2014). A satisfactory unique definition of NUE that can describe this trait across species and environmental conditions is not available (Reich, Aghajanzadeh, and De Kok 2014). Efficiency in general is defined as the achievement of an intended outcome with the lowest possible input. In the particular case of the NUE the input is nutrients, the intended outcome can vary according to the context considered. In an ecological and evolutionary context, the outcome is the optimal and balanced vegetative and reproductive growth suitable to survive and compete in their respective habitat and niche through the use of temporal and spatial use of nutrients. In an agricultural context the intended outcome shifts from the offspring to the desired product yield that can be used as food or for other economic purposes (Reich, Aghajanzadeh, and De Kok 2014). Breeding for the desired product yield creates a detachment of the nutrient use efficiency from its evolutionary and ecological significance. Firstly, in discussing the NUE in an agricultural context, it is important to specify the level at which it is intended, agro-ecosystem, field and plant. The agro-ecosystem is favoured when the NUE is analysed from a holistic perspective. Since the field is the economic unit for farmers, this level is favored for agronomic calculation,. For scientific research purposes the plant level is often used (Reich, Aghajanzadeh, and De Kok 2014). Here we are going to use the definition of NUE at a plant level according to the equation:

$$NU^{E} = NAc^{E} * NUt^{E}$$

where NUE is considered the product of the efficiency of nutrient acquisition (NAcE) by the efficiency with which the nutrient is utilized to produce the desired yield (NUtE) (Weih, Hamnér, and Pourazari 2018; Reich, Aghajanzadeh, and De Kok 2014).

Perennial plants, in contrast with annual plants, recycle elements between consecutive growth cycles. Thus, the definition of NUE provided above has to be widened and adapted

to include different plant life-cycles. The first definitions of NUE for perennial plants attempted to include the amount of nutrient in the plant and the nutrient losses to the ecosystem (Vitousek 1982). The NUE in the context of perennial plants can be defined by extending the equation provided by Berendse and Aerts (1987) for nitrogen and that according to Reich, Aghajanzadeh, and De Kok (2014) can be used for other elements:

$$NU^{E} = NP * MR^{T}$$

In this equation the NUE is split into two components: (1) the instantaneous rate of carbon fixation or biomass production per unit N present in the plant (NP) and (2) a measure of the period in which N can be used for carbon fixation (MRT). From this equation it is possible to see that perennial plants, by having a higher value for MRT provided by the ability to store and recycle nutrients, achieve a higher NUE compared with annual plants in environments with limited nutrient availability (Weih, Asplund, and Bergkvist 2011 ; Dawson, Huggins, and Jones 2008).

The law of the minimum [Sprenger (1839); Reed (1942); Ploeg, Böhm, and Kirkham (1999); Gorban et al. (2011)), also called Law of Liebig, states that the productivity of plants is limited quantitatively and qualitatively by the scarcest resource. The law is not universal, and it is often violated in well-adapted genotype-environment combinations. This exception of the law was defined as the "Law of the Minimum Paradox" (Gorban et al., 2011). A consequence of the Law of The Minimum Paradox is that if we observe the Law of Liebig in artificial systems, then under natural conditions the law will be violated (adaptation will equalize the load of different factors). In contrast, if the law is violated in an artificial system, then it can be assumed that it will be obeyed under natural conditions (adaptation will compensate the violation). According to the Law of the Minimum Paradox, it is possible to hypothesize that when a well-adapted genotype from the wild is grown artificially out of the environment where it is adapted, this restores the validity of

the law of Liebig until new adaptation. When applied to plant nutrition, the law of Liebig states that it is the nutrient available or taken up in a sub-optimal amount that determines the productivity of the plant. Therefore, a balanced supply of all the macronutrients is required. Thus, fertilisation techniques based on few macroelements are valid as long as all the other elements are present to an optimal level (Gorban et al., 2011. Moreover, applying the Law as postulated by Gorban et al. (2011) to nutrition leads to the hypothesis that plants adapted to specific soils and nutrient availability conditions are not limited by the scarcest nutrient in the environment where they are adapted. However, when grown in a different environment their performance will be again limited by the scarcest nutrient. In the short term, the reaction of a plant to stressful conditions depends on the ability of the plant to maintain its characteristics in a different environment. This concept is included in the vaster concept of plasticity.

1.5 Aims and experimental approach

The composition of the biomass used can affect the efficiency of the conversion process used and it is crucial for the profitable replacement of fossil sources with biomass. In the last decades, the breeding of *Miscanthus* for biomass production has mostly focused on the improvement of the biomass yield and only recently optimization of biomass quality in terms of chemical composition and physical structure in relation to specific conversion processes started to receive attention (Weijde et al. 2013; da Costa et al. 2019; Pancaldi and Trindade 2020). Biomass chemical composition and physical structure affect each conversion route in a specific way, thus an unequivocal general definition of the traits for biomass quality cannot be provided. However, quality traits can be identified when a specific end-product and process are defined in the first place. For example, when the quality is defined in relation to the recalcitrance to deconstruction, then the quality can

be seen as the ability to release the components of the cell wall through enzymatic reaction (saccharification) with the minimal input in terms of energy and enzymes. Within the plant the traits determining the quality of biomass for conversion can vary according to the organ considered. Da Costa et al. (2019) found that in *Miscanthus,* lignin has a predominant effect in stems for determining recalcitrance, while in leaf biomass the recalcitrance is mostly related to the nature and ornamentation of matrix glycans. Once the target product/process and the related quality traits have been defined, two approaches can be used to optimise and improve the quality of biomass for conversion (van der Weijde, 2013. The first approach involves the achievement of enhanced biomass quality through breeding for quality traits. The existence of extensive genetic variability in these traits in *Miscanthus* makes plausible the use of breeding to improve biomass quality (Clifton-Brown et al. 2008; Heaton, Dohleman, and Long 2009). The genus *Miscanthus* includes numerous species that could represent a wide genetic base for future breeding programs. However, up to this moment, most of the research in temperate environments has focused on *M*. x *giganteus* and its parents *M. sacchariflorus* and *M. sinensis*.

The second approach involves the improvement of biomass quality through 'on-field quality management practices' such as fertilization and harvest time. Considering that the growth of *Miscanthus* as a bioenergy crop will predominantly happen on marginal lands, the effects of the peculiar characteristics of these environments on the quality of biomass must be considered. Although the impact of abiotic stresses on cell wall has been reported (Le Gall et al. 2015), there is little information about the effect of nutrient limitation on the cell wall characteristics and the quality of biomass for conversion. Le Gall et al. (2015) provides a comprehensive review of the effects of abiotic stresses on the cell wall. The author concluded that although there is general agreement on the effect of abiotic stresses

on cell wall architecture, the amount of the data available differs depending on the stress considered. While there is extensive knowledge about cell wall response to drought or cold stress, information concerning other stresses, such as flooding and air pollutants is still limited and is mostly more concerned with transcriptome and proteome analysis than the detailed chemical composition of the cell wall in response to stress (Le Gall et al. 2015). Weijde et al. (2016) reported the effects of drought stress on biomass quality traits in *Miscanthus*. In particular they found that drought stress resulted in a decrease in the amount of cellulose content in the cell wall, but, interestingly, also in a decrease in recalcitrance to deconstruction. In addition, they found that the effect was affected by the genotype considered. Overall, the study suggested the availability of a genetic variability in *Miscanthus* for future breeding for these traits under drought stress conditions. Amongst the abiotic stresses, and in spite of its metabolic and environmental importance, the effect of nutrient limitation on cell wall has received very little attention. For example, is not mentioned in the review by Le Gall et al. (2015). Some information about the effect of nutrient limitation on cell wall has been provided in recent years. Ogden et al. (2018) reviewed the effects of nutrient stress on cell wall structure and composition in the roots, highlighting the remodeling of the cell wall that allow this organ to optimise nutrient uptake under nutrient limiting conditions. A gene co-expression analysis Landi and Esposito (2017) reported that genes related to nitrogen uptake and to cell wall structure were co-expressed in *Arabidopsis thaliana*. There is even less knowledge on the effect of nutrient stress on type II cell wall of grasses. The only study where this was studied is the investigation by Costa et al. (2018), who studied the effect of a general low nutrient scenario in *Miscanthus*. However, the effect of the limitation of specific nutrients on cell wall composition in commelinids and specifically in *Miscanthus* has not been reported to date. In addition, it is well known that artificially created hybrids of *Miscanthus* have

excellent performances in terms of yields. Still, their biomass quality for conversion has not been explored. The experimental work described here aimed to test effects of nutrient limitation on plant organ in relation to biomass related traits in species of the genus *Miscanthus* and their hybrids. This hypothesis was tested by investigating the extent of variability of cell wall characteristics in genotypes of the genus *Miscanthus* and some of their hybrids and to assess the effects of nutrient stress on their cell wall characteristics. The plant material was obtained from a large germplasm collection maintained in a field trial at IBERS (Aberystwyth) including a large genetic and phenotypic variability. Fig. 1.6 shows a picture of the field in 2017 with plants in their mature growth stage.

Figure 1-6 Picture of the field experiment (ABR33) maintained at IBERS.
Biomass material used in the experiments described in this thesis was collected from this field trial. The genotypes were selected to include most of the genetic variance available in the germplasm collection of IBERS. The work of this thesis was based on genotypes from this field trial.

2 Germplasm from the wild, a source of variability

2.1 Introduction

Miscanthus is a genus of C4 tropical perennial rhizomatous grasses that have the potential to become a specialised biomass crop (Lewandowski, 2003). This process involves domestication similar to practices carried out with other species of monocotyledons in prehistoric times. The aim is to harness the genetic variability underlying the morphological variation between the different species of the genus and genotypes of the same species (Yan et al., 2012) to create varieties suitable for the production of biomass on marginal lands that can be used as a feedstock for biorefining. The main target of breeding programs should include the improvement of the biomass traits that maximise the efficiency of the conversion process considered (quality of biomass for conversion) as well as the ability of the plant to maintain these traits under the stressful conditions expected in marginal lands (Da Costa et al., 2019; Pancaldi and Trindade, 2020) along with the maintenance of high yields of biomass.

The breeding process of new varieties starts from the collection of germplasm in the wild. Subsequently, the germplasm with desirable characteristics is selected to be used in breeding programs (Huang et al., 2018). In this latter phase, to obtain varieties able to produce high yields of biomass with high quality for conversion as defined above, under stress conditions, the genetic and phenologic plasticity of biomass quality traits affecting each conversion process should be considered. Useful information on stress resilience and biomass quality can be provided by the study of the environment to which a species or a genotype is adapted in the wild. *Miscanthus* includes 16 confirmed species (The Plant List, 2013). Although the genus *Miscanthus* has a broad geographical distribution, its species are characterised by a distinctive geographical distribution, and each of them can

be found under typical climatic conditions (Li et al., 2016). In this regard, it is essential to remember that, as in other wild species, the distribution of *Miscanthus* spp. is driven by inter-glacial warming patterns and geographical barriers that can result in populations that are variable in terms of adaptation to different environments (Clark et al., 2014).

It has been shown that the geographical distribution of a species can inform on its vulnerability to climate change (Bernardo, 2014; Sunday et al., 2014). Adaptation of a species to nonuniform climatic conditions indicates its plasticity and consequently resilience to abiotic stress and climate change (Janzen, 1967). Both inter- and intra-specific variability in *Miscanthus* have been well documented (Clifton-Brown et al., 2001; Liu et al., 2013). Inter and intra-specific variability are the result of a process of natural adaptation to environmental conditions. This process resulted in different mechanisms of adaptation and different degrees of resilience (Yan et al., 2012).

This is an exciting fact because the presence of interaction between genotype variation and environmental variation (GxE) is evidence for genetic variation in phenotypic plasticity. In crops, there is a general lack of understanding of physiological and genetic factors influencing GxE and relationships between genotype plasticity and productivity (Aspinwall et al., 2015). Although the modern breeding of *Miscanthus* aims, amongst others, to create new varieties with improved cell wall characteristics for conversion into chemicals and energy, little information is available about the inter and intra-specific extent of variation of these traits. To determine the extent of environmental differences experienced by *Miscanthus* genotypes when grown in European environments, we investigated the environmental characteristics in terms of climate and soil in the areas where *Miscanthus* is native and compared them with conditions in a European environment. Specifically, we used genotypes from a germplasm collection including ~1000 genotypes maintained in Aberystwyth, Wales. We compared conditions of the

climate and soil in the areas where the genotypes were collected in the wild with the ones in Wales. The aim was to select a group of genotypes representing the genetic variability of ABR 33.

2.2 Material and methods

2.2.1 Description of the *Miscanthus* population

The *Miscanthus* population used for the selection is part of an association mapping field trial, identified as ABR33, maintained at the Institute of Biological Environmental and Rural Sciences (IBERS) in Gogerddan, Aberystwyth (position of row 1 column 1: Lat 52.432472 Long -4.020582 Altitude 27.6 above sea level). The field was established in 2012 and included 1,040 plants replicated three times. The genotypes included in ABR33 were selected to gather a wide range of the genetic variability of the germplasm collection available at IBERS. When this experiment began, there were 995 plants alive and identifiable in ABR33. Characteristics of the soil and climatic conditions of the site were obtained in the same way as for the locations of origin of the genotypes described in the "Data sourcing" section below.

2.2.2 Data sourcing

2.2.2.1 Genotype details

Based on data contained in the internal IBERS GenBank and phenotyping platform, MScan (Huang et al., 2018), a complete list of the accessions included in ABR33 was compiled. For accessions originating from propagation material collected in the wild the geographical coordinates of the collection point, determined using the Global Positioning System (GPS), were downloaded in a comma-separated file format that can be used on different platforms.

2.2.2.2 Species distribution in the wild

The distribution of the species in the wild was mapped using data from the Global Biodiversity Information Facility (GBIF). The database of GBIF includes data of species observations with their exact coordinates from different sources (Wieczorek et al., 2012). The number of species observations per geographic point were counted using the function stats::count() in the R programming environment. The distribution map was created using the package ggmap in R programming environment.

2.2.2.3 Climatic data

Climate conditions, which include the variables reported in Table 1, were retrieved for all the genotypes for which coordinates of origin were available using R (R Core Team, 2018)according to (Hijmans and Elith, 2017) from WorldClim database v.1.4. The dataset was downloaded in comma-separated format. The information gathered included temperature related and precipitation related climatic variables.

2.2.2.4 Soil data

Characteristics of the soil for each location were obtained from the Harmonized World Soil Database (HWSD) (Fao/Iiasa/Isric/Isscas/Jrc, 2009). The data were downloaded in Microsoft proprietary format and converted to a SQL 3 format using MS Access to Sqlite3 Converter software freely available for download on GitHub (https://github.com/sanandrea/mdb2sq3). The information was downloaded and organised in a comma-separated file. The database contains information for the chemical and physical properties of the topsoil (TS, 0 – 100 cm from ground level) and the subsoil (SS, >100cm from ground level). In the database, the data was organised in soil map units. A soil map unit is a collection of areas defined and named the same in terms of their soil components, miscellaneous areas, or both. For each soil unit, the contribution to the percentage of different soil classes is reported. For each soil class, the physical and

chemical characteristics of top and subsoil are indicated. To obtain a single value able to summarise the features of the soil, the weighted arithmetic mean was calculated. A query was performed on the same databases using the coordinates of the location of ABR 33.

2.2.3 Data analysis

2.2.3.1 Software used

The data analysis was performed using the R programming language (R Core Team, 2018). Along with the functions from 'stats' package, additional packages were required to complete the elaboration of the data described below. In particular, the following packages were used:

- plyr (Wickham, 2011) for tidying up, summarise and combine datasets
- ggfortify (Tang et al., 2016) and ggplot2 (Kahle and Wickham, 2013) to generate plots from multifactor analyses
- missMDA (Josse and Husson, 2016) for handling the missing values prior multivariate analysis and impute datasets

2.2.3.2 Construction of a complete dataset

The datasets were homogenised using the variable containing the name of the genotype as a unique identifier and merged using merge () function. The missing values in the dataset were predicted using the iterative PCA algorithm (Josse and Husson, 2012) by a Bayesian method (Verbanck et al., 2013). The algorithm works only if the variables are linear. Therefore, non-linear variables were log-transformed before imputation.

2.3 Results

2.3.1 Geographical details of the area of origin of the genotypes in ABR33

Complete climatic and soil condition information was retrieved for the location of a collection of genotypes maintained in the ABR33 association field trial. In particular, it was possible to obtain geographical coordinates of origin for 599 out of 995 accessions

(alive at the moment of the data collection) included in ABR33. After checking for annotation errors, the geographical coordinates (Latitude and Longitude in decimal format) for 592 genotypes were confirmed. For each genotype, information regarding its genetic group was obtained from Dr Gancho Slavov (unpublished data). The genotypes in the population from ABR33 are assigned to 8 genetic groups according to a discriminant analysis of principal components (DAPC) of a genetic characterization based on genetic markers

The eight groups are:

- *M. sinensis* from South Japan
- *M. sinensis* used in the EMI/PRI project
- M.sinensis from North japan
- M. sinesis from Taiwan
- *M. floridulus* from Taiwan
- *M. sacchariflorus* / M. robustus
- Hybrids naturally occurring
- M. lutatioriparius

Miscanthus species have been reported to occupy specific geographic areas. The distribution of *Miscanthus* species, according to the data downloaded from GBIF, is presented in Figure 2-1. The figure shows there are 10 species reported in East Asia. Amongst them, *M. fuscus* is the one occupying the largest area. It is distributed between India, Laos and Cambodia. Its area of distribution partly overlaps with one of *M. nudipes* and *M. nepalensis*. The distribution of the species in the wild was compared with the coordinates of the locations were the genotypes in ABR33 were collected. Only 5 of them are represented in ABR33 (*M. sacchariflorus*, *M. sinensis*, *M. floridulus*, *M.*

lutarioriparius and *M. x giganteus*). Genotypes identified as *Miscanthus* sinesis and *Miscanthus* sacchariflorus/robustus in ABR33 were collected out of the area where these species are reported in the wild according to the GBIF (Figure 2-2 A and B). From Figure 2-2 C and D, it is possible to notice that *Miscanthus* floridulus and *Miscanthus* lutarioriparius have a vast area of distribution in the wild in the South of China. However, genotypes of these species in ABR33 were collected from the core of the distribution area. The genetic groups distributed on a different range of latitude, longitude and altitude in the area of origin (Figure 2-3, Figure 2-4 and Figure 2-5). Most of the genotypes, with the exclusion of those belonging to *M. floridulus* and *M. sinensis* from Taiwan, are distributed in a latitudinal range between 25- and 42-degrees North of the equator. Genotypes belonging to *M. floridulus*, live in a narrower range of latitudes, between 18- and 25-degrees North of the equator. Most of the genotypes in the population under study came from a longitudinal range between 120 and 140 degrees East of Greenwich, irrespective of the genetic group they belonged to (Figure 2-4). Equally, almost all the genotypes came from areas in altitudes between 0 and 1000 meters above sea level (m a.s.l.). Genotypes living above this limit belong to *M. sinensis* and *M. floridulus* groups. Two genotypes from Taiwan, belonging to *M. floridulus* and *M. sinensis*, respectively, were collected at altitudes close to 3000 m a.s.l (Figure 2-5).

Taken together, the results of these observations pointed to differences in extent and shape between the area of collection of the germplasm and the area where the genotypes were reported in the wild. Besides, comparing the geographical location of the collections with the area of distribution of each species (Fig. 2.2) the germplasm was mostly collected from the core area of the distribution region.

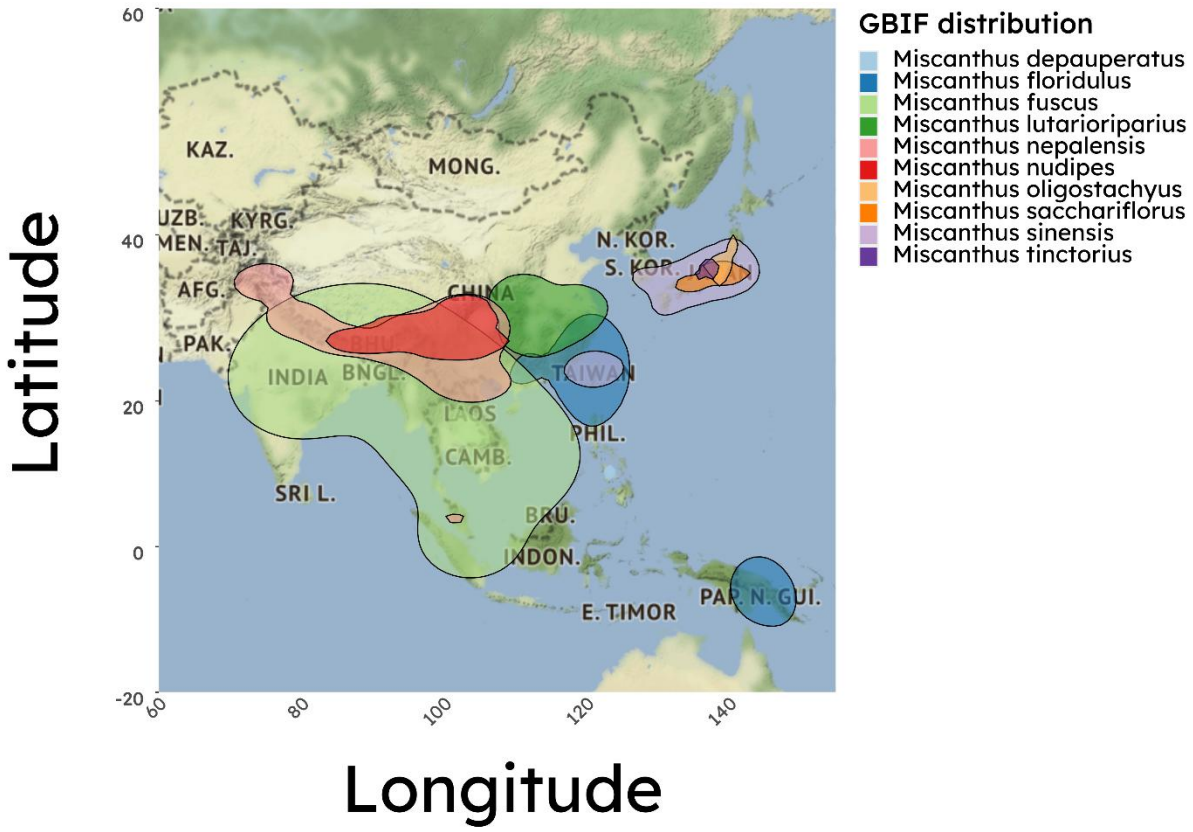


Figure 2-1 Area of distribution of *Miscanthus* species in the wild.
The distribution of *Miscanthus* species is presented using data from GBIF. Amount of observations for a geographic point were counted. The shaded areas represent the result of kernel density estimation. The plot was realised in an R environment using package ggmap.

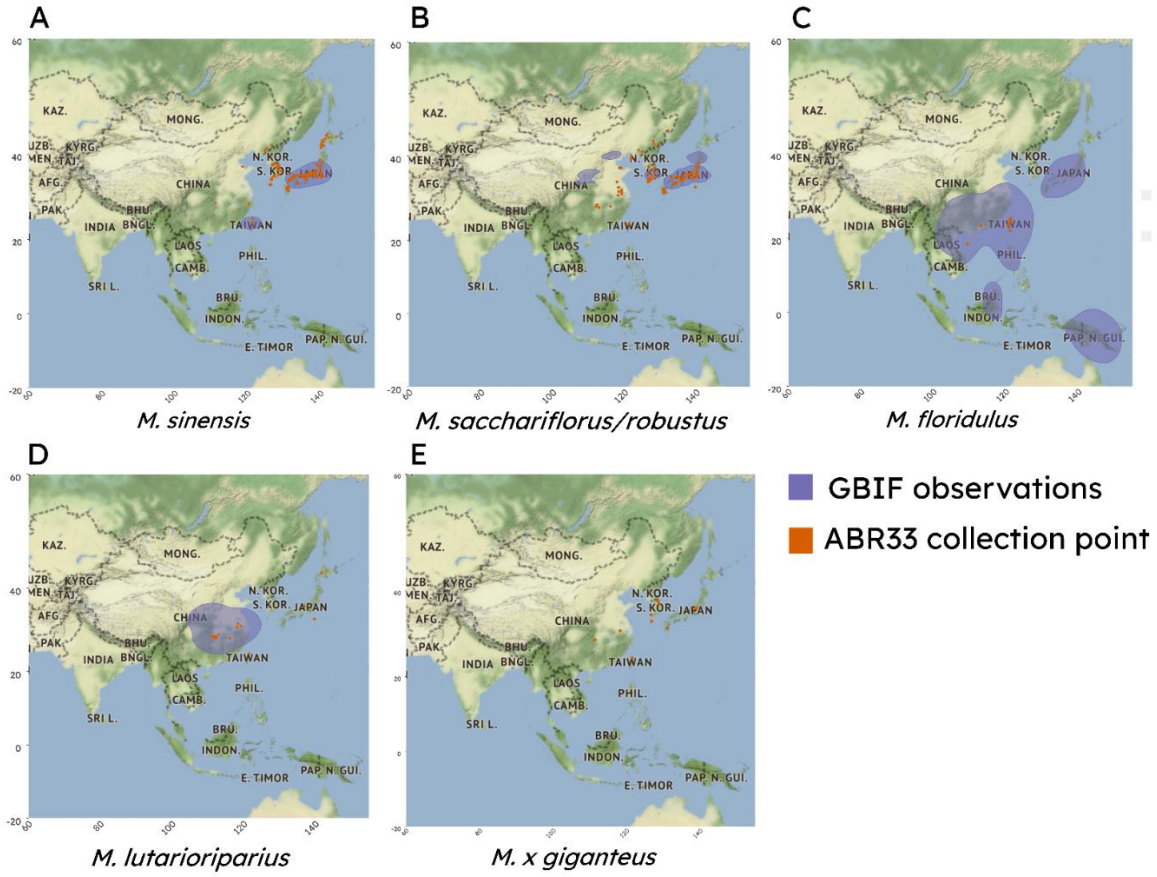


Figure 2-2 Geographical distribution in the wild of *Miscanthus* species and collection points of the *Miscanthus* germplasm.
The geographical distribution of a species is given by the natural arrangement and apportionment of the various *Miscanthus* species in the different regions and localities. Using the data from GBIF and from MScan, the points of collection of genotypes in ABR33 were compared to the area of distribution of the species in the wild. The plot was realised in an R environment using package ggmap.

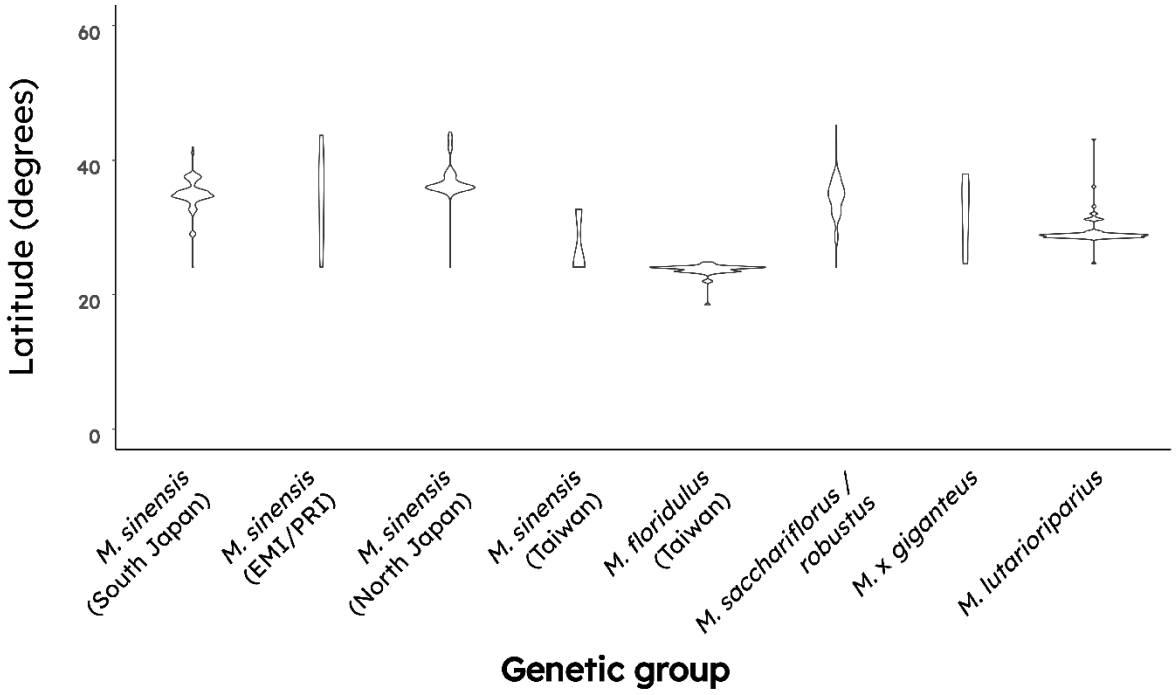


Figure 2-3 Range of distribution of the genotypes in latitude.
Range of latitude in which genotypes from the different genetic groups were collected. Values on the y-axis represent the value of latitude in degrees where each genotype in ABR33 was obtained. The violin shape describes the kernel density estimation to show the distribution shape of the data

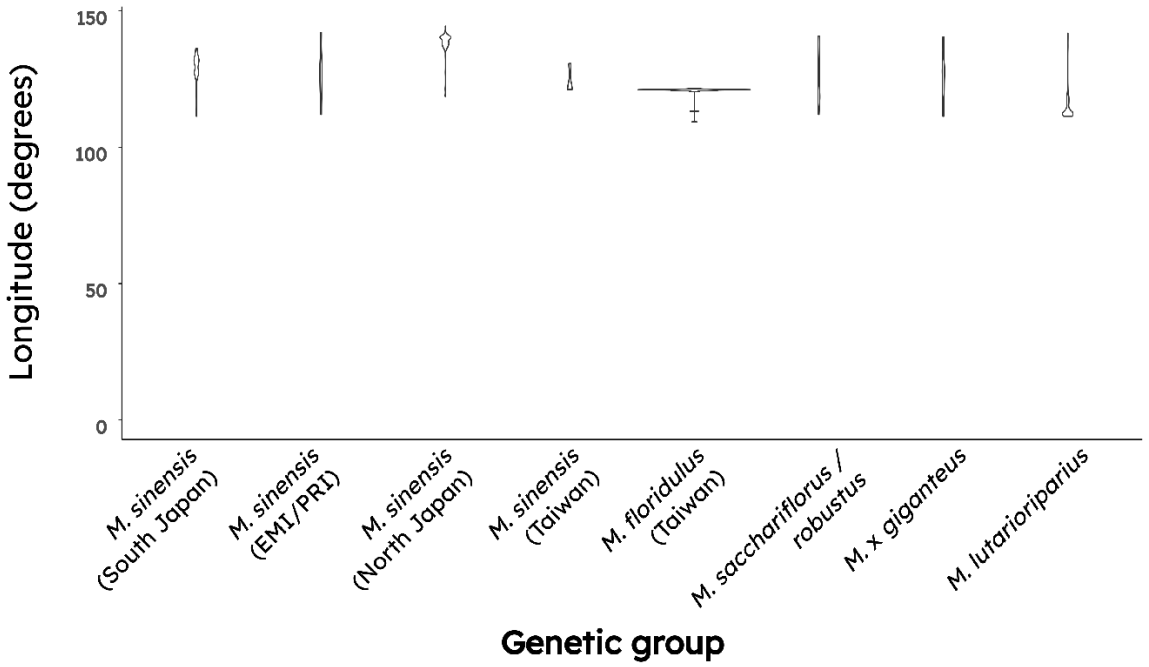


Figure 2-4 Range of distribution in longitude.
Distribution of the genotypes in the range of distribution in degrees longitude of the respective genetic group. Values on the y-axis represent the value of the latitude in degrees where each genotype in ABR33 was collected. The violin shape describes the kernel density estimation to show the distribution shape of the data.

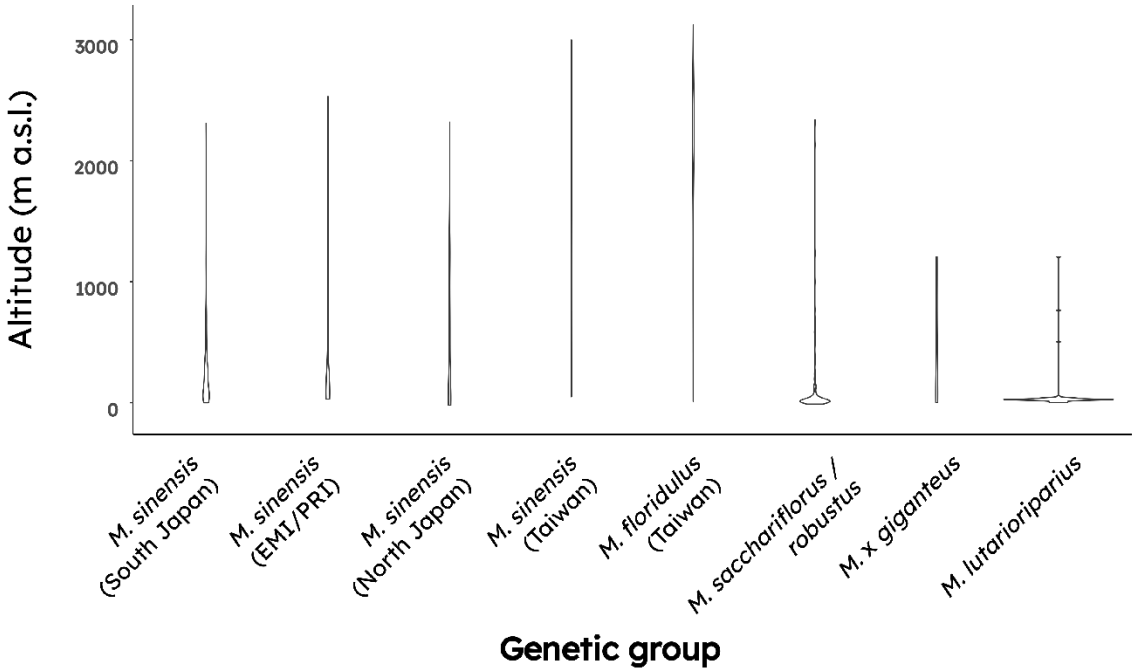


Figure 2-5 Range of distribution in altitude.
Distribution of the genotypes in ABR33 field in the range of altitudes where they were initially collected. Values on the y-axis represent the value of the altitude in meters above sea level (m a.s.l.) where each genotype in ABR33 was collected. The violin shape describes the kernel density estimation to show the distribution shape of the data.

2.3.2 Climatic and soil variables obtained for each genotype

For each location where genotypes in the population under study were collected, soil and climatic variables were downloaded using the database mentioned before. The climate variables and their description are listed in Table 2-1 and the soil variables and their description in Table 2-2. The quality of the data was checked for accuracy and consistency and data points that contained values not acceptable were removed. In particular, for two genotypes, the coordinates were annotated as Latitude 0 and Longitude 0, this was attributed to an error in annotation, and the two data points were removed from the dataset. The list also included some genotypes from Germany, these genotypes are not native of the area but imported from Asia, and they were consequently excluded from the dataset as well.

Climatic and soil data for each genotype were joined together in a single dataset. Shapiro-Wilkinson test for normality of the dataset showed that none of the variables was normally distributed and that the distribution was skewed. The reason for the skewness of the variables is not surprising and can be attributed to the presence of *Miscanthus* species in some specific soil-climate conditions. The dataset contained some missing values, especially when SS data were regarded. Different techniques are mentioned in the literature to fill in the gaps and be able to perform a principal component analysis (Little et al., 2002; Josse and Husson, 2012). Here, each missing value was replaced with the median value for each variable across all the location points.

Soil and climate variables in the area of origin correlate

The complete dataset containing information about the soil and climatic conditions of the area where each accession was collected, and its genetic identity, was used to perform a correlation analysis of all the variables using the Pearson method with Bonferroni correction to describe the typical environmental characteristics of the areas where each

accession lives in the wild. For example, the pH of the soil (S_PH and T_PH) had a negative correlation with precipitation variables and not with the temperature variables, which indicates acid soils under high precipitation conditions. The information is presented by a correlation matrix with the significance of the correlation and the R squared graphically represented (Figure 2-6).

It is interesting, although not surprising, to notice that there was a strong negative correlation between the total annual precipitation (PTA) in each location and the pH in the topsoil (T_PH) and subsoil (S_PH). However, the regression coefficient was no higher for T_PH compared to S_PH.

The topsoil and subsoil sodicity (T_ES, S_ES), the topsoil and subsoil salinity (T_EC, S_EC) and the subsoil gypsum (S_CO) were positively correlated to the annual seasonality of precipitation. T_EC and S_EC are the electrical conductivity measured in the topsoil and subsoil, respectively. They are an estimation of the salt content. T_ES and S_ES are the exchangeable sodium percentage (ESP) in the top and subsoil. The ESP was used to indicate levels of sodium in soils. The value was obtained as the ratio of Na in the CEC according to Equation 2.1.

Equation 2.1

$$ES^{P} = \frac{Na \cdot 10^{0}}{CE^{C}_{soi}} l$$

The S_CO is describing the calcium sulphate (gypsum) content in the subsoil. Gypsum has been reported to occur occasionally in soils, particularly in the driest areas of the globe. The seasonality of annual precipitation (PSA) is determined as the coefficient of variation of the average yearly rainfall (PAA).

Taken together, these results describe the variability of the climatic condition in which species of *Miscanthus* can live. Also, they provide information that can be exploited to assist future germplasm selections.

Table 2-1 List of climate variables used

List of the variables that were used in this analysis to describe precipitation, temperature and their variation from data collected in the last 50 years. Each variable is identified by a variable name (Variable), an ID that will be used in figures, the metric unit used to measure the variable and a description of the variable.

Climate variates			
Variable	**ID**	**Unit**	**Description**
TempAverAnn	TAA	°C	Annual average temperature
TempRangDay	TRD	°C	The annual average of daily range of temperature variation1
TempIsotAnn	TIA	%	Isothermality. The highest value of the percentage of the annual temperature variation perceived in a day.2
TempSeasAnn	TSA	%	Coefficient of variation of TAA
TempMaxAnn	TMA	°C	Maximum annual temperature
TempMinAnn	TmA	°C	Minimal annual temperature
TempRangAnn	TRA	°C	Annual range of variation of temperature
TempAverWeQ	TWeQ	°C	The average temperature in the wettest quarter of the year

1 Mean Diurnal Range (Mean of monthly (max temp - min temp))

2 (TRD/TRA)*100

Climate variates			
Variable	**ID**	**Unit**	**Description**
TempAverDrQ	TDQ	°C	The average temperature in the driest quarter of the year
TempAverWrmQ	TWaQ	°C	The average temperature in the warmest quarter of the year
TempAverColQ	TCQ	°C	The average temperature in the coldest quarter of the year
PrecTotAnn	PTA	mm	Total annual precipitation
PrecTotWetM	PWM	mm	Total precipitation in the wettest month
PrecTotDryM	PDM	mm	Total precipitation in the driest month
PrecSeasAnn	PSA	mm	Annual seasonality of precipitation. Coefficient of variation of PTA.
PrecTotWetQ	PWeQ	mm	Total precipitation in the wettest quarter
PrecTotDryQ	PDQ	mm	Total precipitation in the driest quarter
PrecTotWarQ	PWaQ	mm	Total precipitation in the warmest quarter
PrecTotColQ	PCQ	mm	Total precipitation in the coldest quarter

Table 2-2 List of the soil variables retrieved from the Harmonised World Soil Database (HWSD)

List of the variables that were used in this analysis to describe the physical and chemical characteristics of each soil map unit. The same set of variables was obtained for the topsoil and subsoil layers. Topsoil is represented by the first 100cm of the soil profile. Subsoil describes the soil below 100 cm of the soil profile. Variables from the topsoil have an ID starting with T and variables form the subsoil start with S. For each variable the table indicates name (Variable) and a short identifier used in figures (ID), the metric unit in which the variable was measured and a brief description of the variable.

Soil variates			
Variables	**ID**	**Unit**	**Description**
T_GRAVEL	T_GR	%vol	Topsoil gravel content
T_SAND	T_SA	%wt	Topsoil sand content
T_SILT	T_SI	%wt	Topsoil silt content
T_CLAY	T_CL	%wt	Topsoil clay content
T_REF_BULK_DENSITY	T_RB	kg/dm3	Topsoil reference bulk density
T_BULK_DENSITY	T_BD	kg/dm3	Topsoil bulk density
T_OC	T_OC	%wt	Topsoil organic carbon content
T_PH_H2O	T_PH	-log(H+)	pH in the water extract of the topsil
T_CEC_CLAY	T_CE	cmol/Kg	Topsoil cation exchange capacity (CEC) in clay fraction
T_CEC_SOIL	T_CS	cmol/Kg	Topsoil cation exchange capacity (CEC)
T_BS	T_BS	%	Topsoil base saturation
T_TEB	T_TE	cmol/Kg	Topsoil total exchangeable bases
T_CaCO3	T_CA	%wt	Calcium carbonate content in the topsoil
T_CaSO4	T_CO	%wt	Calcium sulphate (gypsum) in the topsoil
T_ESP	T_ES	%	Topsoil sodicity
T_ECE	T_EC	dS/m	Topsoil salinity

Soil variates			
Variables	**ID**	**Unit**	**Description**
S_GRAVEL	S_GR	%vol	Subsoil gravel content
S_SAND	S_SA	%wt	Subsoil sand content
S_SILT	S_SI	%wt	Subsoil silt content
S_CLAY	S_CL	%wt	Subsoil clay content
S_REF_BULK_DENSITY	S_RD	kg/dm3	Subsoil reference bulk density
S_BULK_DENSITY	S_BD	kg/dm3	Subsoil bulk density
S_OC	S_OC	%wt	Subsoil organic carbon content
S_PH_H2O	S_PH	-log(H+)	pH in the water extract of the topsil
S_CEC_CLAY	S_CE	cmol/Kg	Subsoil cation exchange capacity (CEC) in clay fraction
S_CEC_SOIL	S_CS	cmol/Kg	Subsoil cation exchange capacity (CEC)
S_BS	S_BS	%	Subsoil base saturation
S_TEB	S_TE	cmol/Kg	Subsoil total exchangeable bases
S_CACO3	S_CA	%wt	Calcium carbonate content in the topsoil
S_CASO4	S_CO	%wt	Calcium sulphate (gypsum) in the topsoil
S_ESP	S_ES	%	Subsoil sodicity
S_ECE	S_EC	dS/m	Subsoil salinity

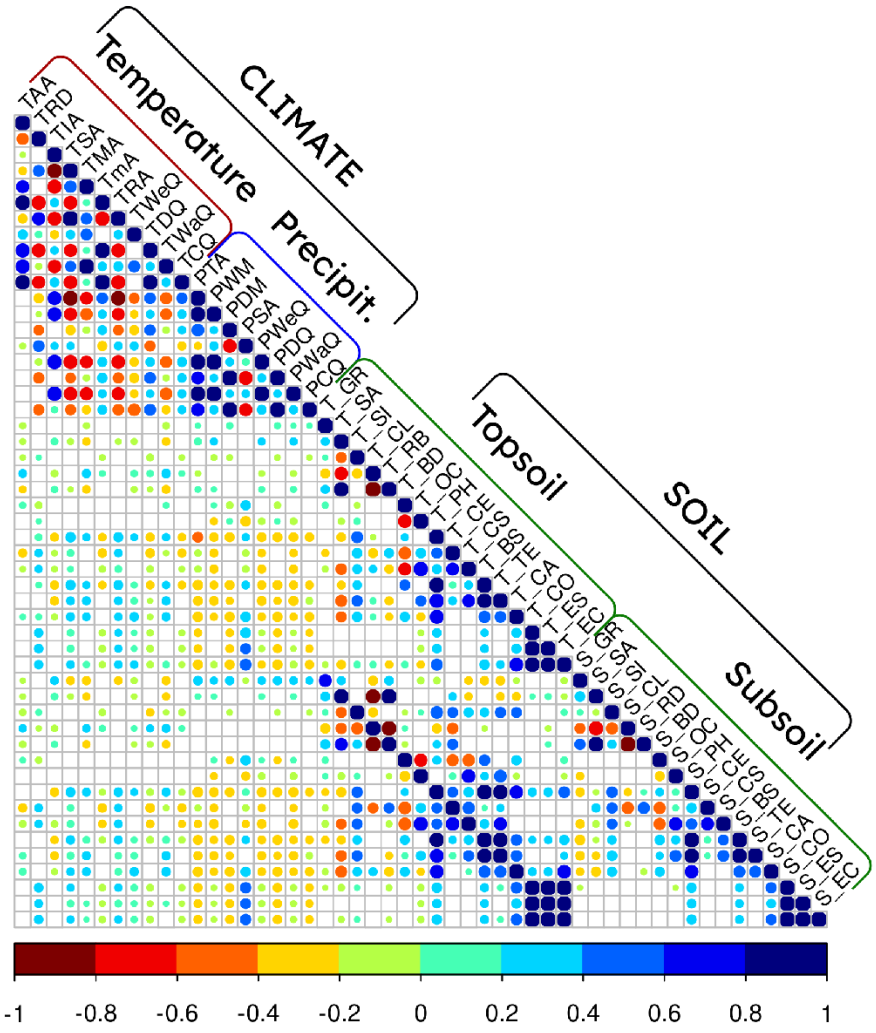


Figure 2-6 Correlation matrix of the variates
Correlation matrix soil and climate information for the 592 genotypes for which coordinates were collected. Significance values of the regression were corrected using a Bonferroni method. Color scale represents the R value of the correlation. Only correlations with $p<0.05$ are shown (blank squares are $p>0.05$.

2.3.3 Principal component analysis (PCA)

To answer the question whether, there was a relation between the genetic group of each accession and the climatic conditions in which they were living, a principal component analysis (PCA) was applied to the dataset containing the pedo-climatic variables. The dataset was scaled and centred before the analysis was performed. The graphical representation of the scores of the PCA allows visualising in which pedo-climatic condition each accession was collected in the wild. The principal component analysis of

the dataset revealed that the first three components were explaining 53.17% of the total variability (Figure 2-7).

Based on the contribution of the variables to the first three principal components, the 20 most contributing variables were identified (Figure 2-8). The variates that had the highest contribution to the first component included PTA, T_PH, S_PH, TSA, TRA, T_BS, S_BS, T_TE, T_EC. These variates all relate to the acidity of the soil and connected climatic conditions. The distribution along the first component could, therefore, be related to the distribution to different soil pHs. The highest contribution to the second component included S_CL, S_RD, T_CL, T_RB, T_CE, S_CE, S_SA, TAA. This group comprises variates relating to the soil texture and the bulk density and the average annual temperature. Finally, the variates that contributed to the third component included T_OC, PSA, T_CS, S_ES, S_CS, S_EC, T_ES, S_CO, T_OC. This is a group of variates that relates mostly to organic soil carbon, salinity and bases content. Taken together, these observations seem to indicate that the three components of the PCA were describing distinct aspects of soil-climatic conditions.

Looking at the plots of the scores from the PCA for every single accession (Figure 2-9), it is possible to see that the climatic and soil condition are distinguishing most of the *M. floridulus* from the other genetic groups. The scores of the genotypes in the different genetic groups in ABR33 are not separated along any of the three dimensions considered. Among the exceptions were two genotypes of M.floridulus (Mb 851 and Mb 855) clustering with the other groups and three genotypes of *M. sacchariflorus* (Mb 1029, Mb 1179, Mb 458) grouping with the *M. floridulus*. These results suggest that *M. floridulus* genotypes are found in climate-soil conditions that are distinct from those where the other genotypes have been collected.

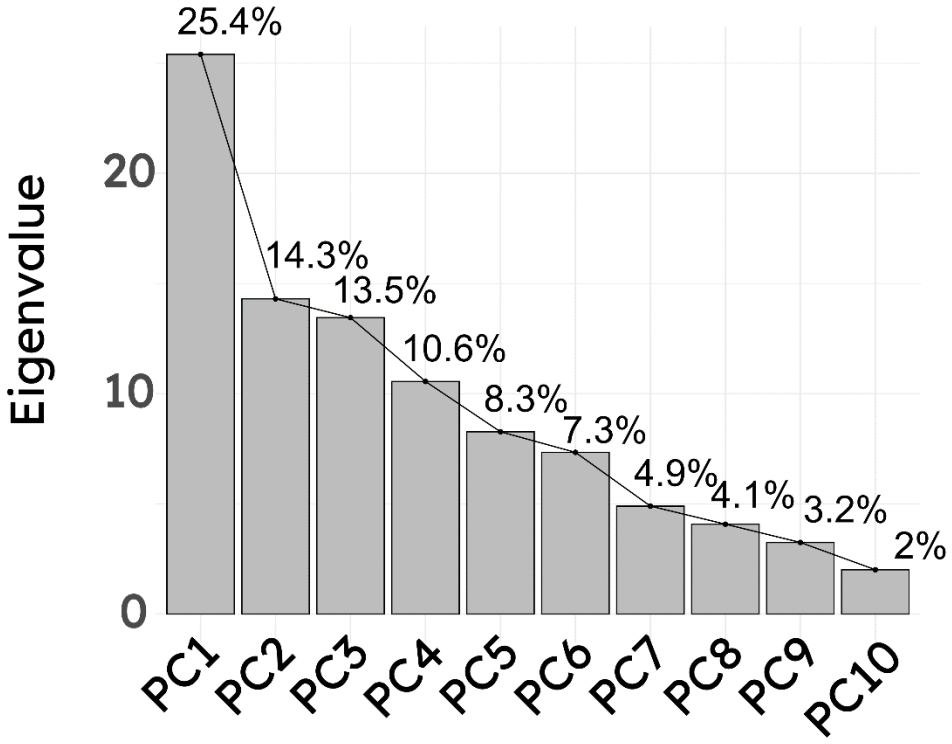


Figure 2-7 Percentage of explained variance by the principal component analysis (PCA).

Screeplot of the principal components of the PCA. Values on the y-axis represent the eigenvalues of each component. An eigenvalue expresses the percentage of variance explained by the respective component. Values on the x-axis represent the components. Only the first 10 components are presented. The first three principal components (PC1, PC2 and PC3) were able to explain 53.2% of the total variance in the dataset.

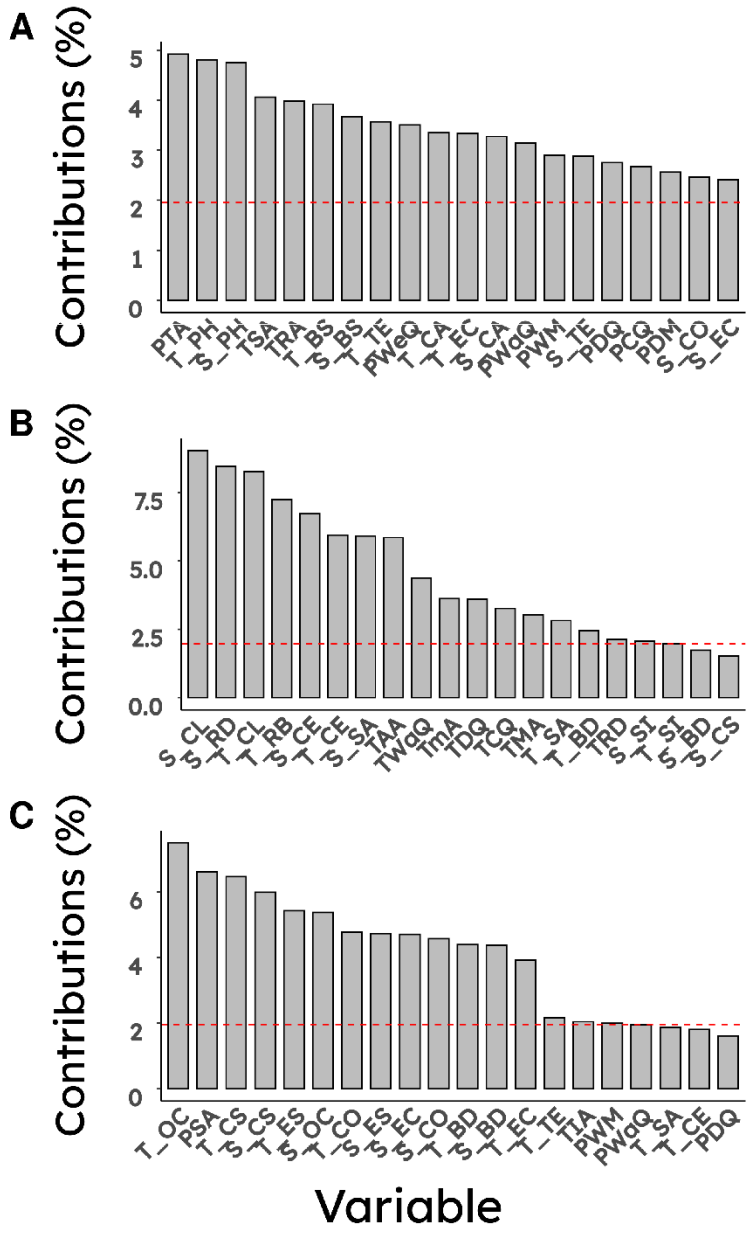


Figure 2-8 Contribution of the variates to the first three components of the principal component analysis (PCA).
Values on the y-axis represent the contributions (in percentage) of the variables to the principal components. Labels on the x-axis indicate the variable considered. Description of the variables can be found in Table 2.1 and table 2.2. The red dashed line on the graph indicates the expected average contribution if the contribution of the variables were uniform.

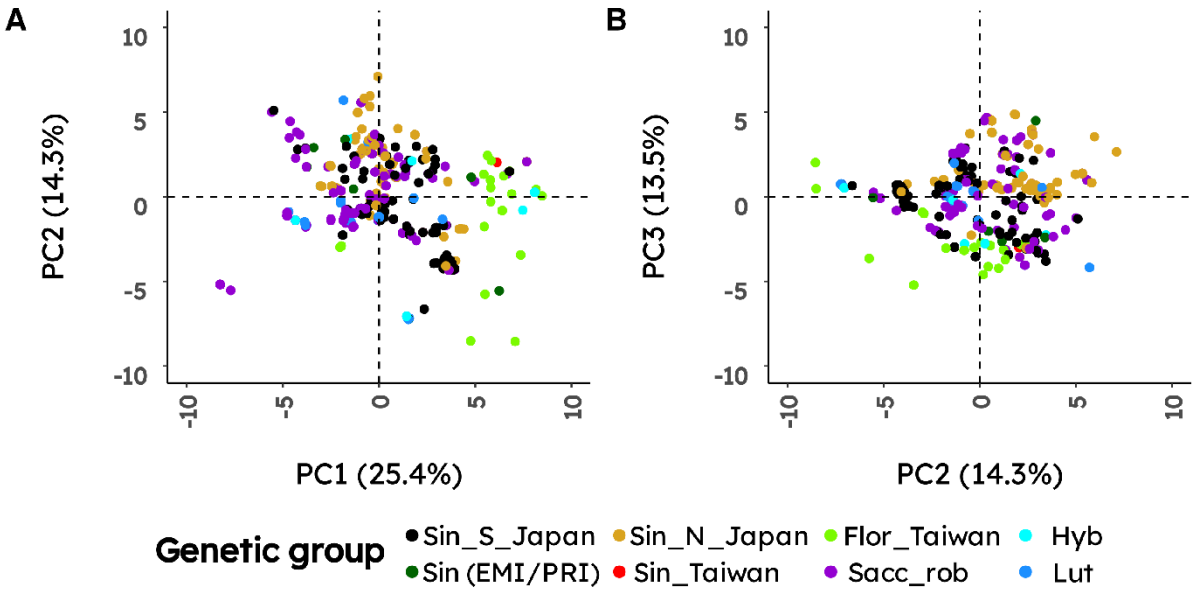


Figure 2-9 Score plot of the first three components of the principal component analysis (PCA).
The first three components of the PCA explained 53.2% of the total variance in the dataset. A. Score plot of the first (PC1) and second (PC2) component of the PCA. B. Score plot of the second (PC2) and second (PC3) component of the PCA. Values on the x and y axes represent the PCA scores of the principal component associated with the axis. Colours represent the genetic group of each genotype. Sin_S_Japan = *M. sinensis* (South Japan), Sin_N_Japan = *M. sinensis* (North Japan), Flor_Taiwan = *M. floridulus* (Taiwan), Hyb = Naturally occurring M. xgiganteus, Sin (EMI/PRI) = *M. sinensis* (used in the European *Miscanthus* Improvement project), Sin_Taiwan = *M. sinensis* (Taiwan), Scc_rob= *M. sacchariflorus*/robustus, Lut = M. lutarioriparius.

2.3.4 Selection of 49 genotypes from ABR 33

The variability detected by the PCA in terms of adaptation to different environmental conditions was used to select a subgroup of genotypes for each species. The PC score of each genotype is assumed to describe the conditions to which it is adapted. At least 3 genotypes were selected for each sub-group. From the population, 44 genotypes were arbitrarily selected in a way that they represented all the genetic groups present in the original population. They were coming from the extreme and average of distribution area in terms of climate and soil condition (Figure 2-10). The selection included one genotype of M. x giganteus (Mb311) previously used as reference material in other

studies (da Costa et al., 2014; da Costa et al., 2017). In addition, 5 artificial hybrids produced at IBERS, having at least one parent included in the initial 44-genotypes group, were added to the selection. The complete list with detailed information of the 49 genotypes selected is shown in Table C. 1.

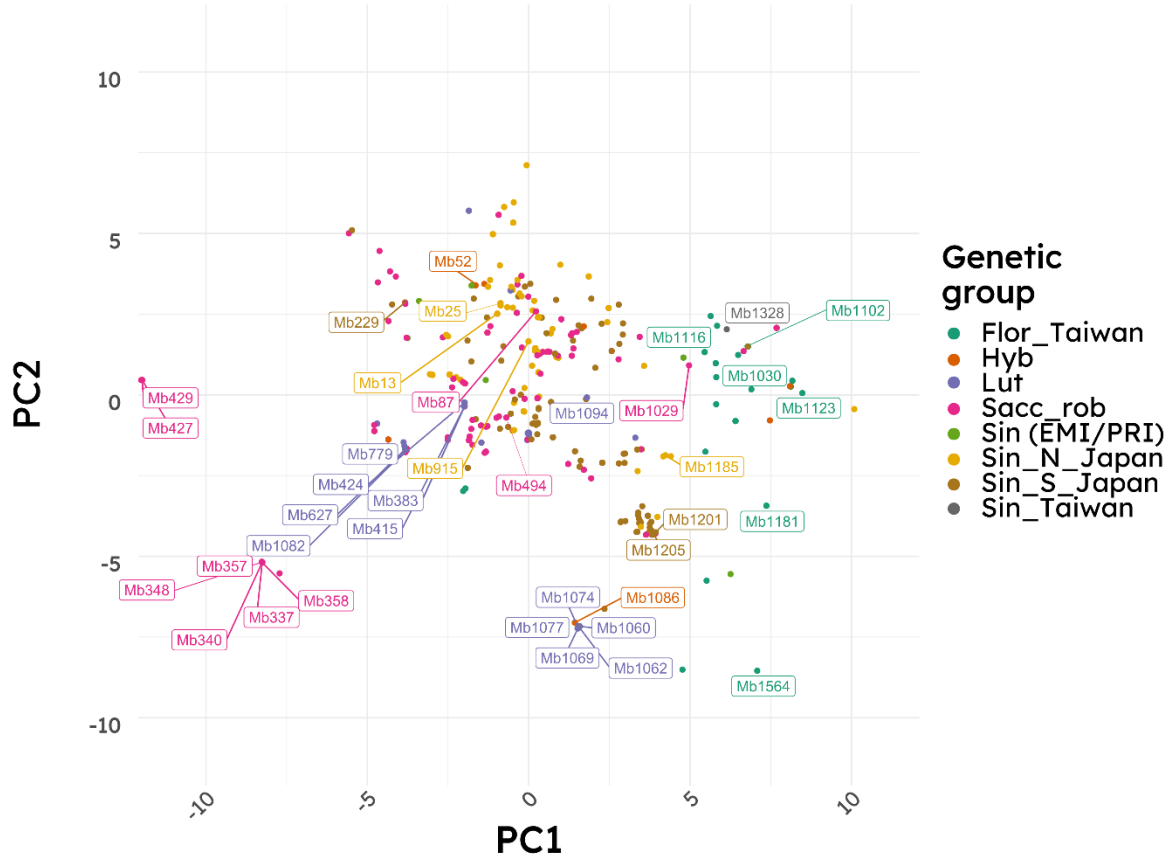


Figure 2-10 Genotype selection based on principal component scores
The PCA scores were used for the selection of 49 genotypes that were representative of the whole population in terms of species and climatic distribution. Values on the x and y axes represent the scores of the PCA on the first (PC1) and second (PC2) components respectively. Labels indicate the name of the genotypes. Colours of the data points indicate the genetic group. Sin_S_Japan = *M. sinensis* (South Japan), Sin_N_Japan = *M. sinensis* (North Japan), Flor_Taiwan = *M. floridulus* (Taiwan), Hyb = Naturally occurring M. xgiganteus, Sin (EMI/PRI) = *M. sinensis* (used in the European *Miscanthus* Improvement project), Sin_Taiwan = *M. sinensis* (Taiwan), Scc_rob= *M. sacchariflorus*/robustus, Lut = M. lutarioriparius. The selection was performed by choosing genotypes that had scores at the extreme and middle point of the range of variation of the range of variation of the scores for each genetic group.

2.4 Discussion

2.4.1 Area of origin of genotypes in the IBERS collection

The collection of *Miscanthus* genotypes maintained at IBERS is one of the most extensive collections of germplasm available in Europe. ABR33 was planted in 2012 to include the highest variability available amongst the IBERS *Miscanthus* collection. The analysis of the geographic coordinates and altitude of the area of origin allowed to characterise the geographical range of origin for the distinct genetic groups represented in the population under study. Firstly, the coordinates of collection of ABR33 accessions were compared with the areas of distribution of *Miscanthus* species in the wild and with the species reported in literature (Hodkinson et al., 2002; Sun et al., 2010; Hodkinson et al., 2015). The comparison provided evidence of the existence of unexplored areas where germplasm can be collected in East-Asia and species under-represented in ABR33. Out of the 16 species of *Miscanthus* reported in the wild, only 5 are included in ABR33 (*M. sacchariflorus*, *M. sinensis*, *M. floridulus*, *M. lutarioriparius* and *M. x giganteus*).

The range of distribution of the species represented in ABR 33 was in agreement with the one reported by the GBIF and in the Sun et al. (2010) that provided one of the most detailed studies of the morphology and geographic distribution of the species of the genus..

All the genotypes in ABR33 of three genetic groups included in *M. sinensis* were distributed over a wide range of latitudes. In literature, *M. sinensis* has been reported to occur in a range of latitudes larger than *M. sacchariflorus* in Japan, Korea, and China (Hager et al., 2014) and Russia (Clark et al., 2016).

M. floridulus is a species distributed in Taiwan from the river banks until 2000m above sea level, where it is the dominant species (Kao et al., 1998). The distribution of

M. floridulus in the wild mirrors its high adaptability to different temperatures (Kao et al., 1998). Observation of plants in ABR33 agrees with these observations in that genotypes belonging to *M. floridulus* are persistent during the year. This was explained by Kao et al. (1998) with the lack of the avoidance mechanisms of cold resistance of low winter temperature when grown at low altitudes. The same study showed that *M. floridulus* is less resistant than *M. transimorrisonensis* to low temperatures.

M. lutarioriparius is a typical species of *Miscanthus* reported in arid and semi-arid areas (Yan et al., 2015).

Some of the species reported by the GBIF and described by Sun et al., 2010, such as *M. fuscus*, *M. nepalensis*, *M. nudipes*, having interesting characteristics for biomass conversion, were not included in the ABR33 collection. *M. fuscus* is distributed over a large area between India, Laos and Cambodia, partly overlapping with the distribution of *M. nudipes* and *M. nepalensis*. Although *M. fuscus* is included in the GBIF database and described since 1882 (Various, 1882), it was not included in the study by Sun et al. (2010). The species has been reported to have excellent characteristics for methane production (Yoo et al., 2017b).

It was observed that genotypes from *M. floridulus* and *M. lutarioriparius* in ABR33 were mostly collected from the central area of their region of distribution. Collecting material from the more marginal areas of the distribution region of the species in the wild could increase the chances of collecting naturally-occurring inter-specific hybrids. Sun et al. (2010) suggested the existence of hybrids between *M. floridulus* and *M. sinensis* in the areas where the two species cohabit. The collection of accessions only from the central area of the region of distribution could have reduced the chances of including such hybrids in the European germplasm collections. *M. nudipes* an endemic species from Himalaya-Hengduan Mountains has two distinct populations

and has been reported as an interesting species for breeding programmes because of its high tolerance to frost and drought (Ma et al., 2015). Although mostly distributed in the same area as *M. nudipes*, there is lack of information about *M. nepalensis*.

To increase the genetic variability available for the breeding of *Miscanthus*, a further collection of germplasm should ideally be carried in the areas unexplored where *Miscanthus* species have been reported and include species under-represented in *Miscanthus*. The ability to produce *Miscanthus* hybrids could benefit from the extensive collection of germplasm in the areas where more species overlap, and the occurrence of hybrids is more probable.

2.4.2 Correlations between environment variates in the collection areas

The variates describing climate and soil conditions in the geographic points where the genotypes were initially collected were tested for association. Some of the strong significant correlations reported in the results such as soil pH and precipitation and rainfall seasonality and salinity/sodicity, show some features typical of soils under certain climatic conditions.

There was a strong negative correlation between topsoil pH and the total amount of precipitation that an area is receiving per year. It has been suggested that rainfall can be associated with nutrient availability in the soil. Quinto-Mosquera and Moreno-Hurtado (2016) found that the total nitrogen available in a natural soil increased with higher precipitation, from 2,000 mm yr-1 to 8,000 mm yr-1 but tended to decrease above this limit. Furthermore, it is known that soils in high rainfall and high-temperature conditions are often characterised by low pH (Ji et al., 2014). Slessarev et al. (2016) demonstrated that climate is responsible for the soil pH at a global scale. They showed that there is a transition from high to low soil pH under conditions where mean annual precipitation begins to exceed mean yearly potential evapotranspiration.

The pH, by influencing the ionization of salts in the soil, can affect the availability of nutrients in the ground. Thus, it could be hypothesised that genotypes living in these specific conditions could have developed adaptation mechanisms that allow extraction of nutrients from soil under low nutrient availability. In the model species *Arabidopsis thaliana*, it was shown that ecotypes adapted to different availabilities of nitrogen in the soil were differently able to cope with nitrogen stress (Ikram et al., 2012). Besides genetic marker analysis of soil ecotypes of A. lyrate displaying significant genotype - by - treatment responses for seed germination pointed to loci involved in ion transport as the main targets of adaptive genetic divergence (Guggisberg et al., 2018).

The strong positive correlation between the seasonality of the precipitation and topsoil and subsoil sodicity (T_ES, S_ES), topsoil and subsoil salinity (T_EC, S_EC) and subsoil gypsum (S_CO) in the subsoil and topsoil are not surprising. Rainfall usually contributes to desalinization of soils (Sheng and Xiuling, 2004). Soils exposed to high range of precipitation variation can develop high values of salinity with a higher percentage of the salts being sodium (van der Zee et al., 2014). Gypsum accumulation in the topsoil is typical in arid climatic conditions. In these conditions, the rainfall seasonality is usually very low.

It could be inferred from these results that genotypes living in areas with a high range of variation in precipitation are also adapted to salinity and sodicity of the soil. The selection for traits related to resistance to salinity is of particular interest for *Miscanthus*. Breeding programs are including the development of varieties of *Miscanthus* resistant to salinity (Stavridou et al., 2016). The information about the

salinity and sodicity of the soils where the genotypes come from could be used to select germplasm able to better withstand salinity.

2.4.3 Distribution of the genetic groups in the climatic areas

A PCA was carried out to test if there was a relation between the genetic group and the climatic conditions where each genotype was collected. The results revealed that the group of *M. floridulus* is grouping separately from all the others genotypes, indicating that this group occupies a distinct ecological area (Figure 2-9).

However, there were some exceptions where genotypes of M.floridulus clustered with the other groups and three genotypes of *M. sacchariflorus* (Mb 1029, Mb 1179, Mb 458) grouped with *M. floridulus*. These exceptions should be further investigated. Experiments using genotypes of a genetic group that cluster with those of another group could help with studies aimed to understand the biological basis of the adaptation to specific climates and soil conditions.

2.4.4 Selection of 49 genotypes of *Miscanthus* spp.

The selection of the genotypes was performed semi-arbitrarily. However, the aim was to include at least three genotypes for each genetic group. Genotypes from the same genetic group were selected according to their first and second component scores upon PCA. It was assumed that the scores of the PCA were summarising the soil-climate conditions of the collection sites. Given the lack of detailed information about the biomass of the single genotypes in ABR33 and of carrying a high throughput characterization, the arbitrary selection was considered a suitable approach to gather the highest range of variability in ABR33. Genotypes were identified by the UID assigned on MScan. To simplify the identification of the genotypes, the code was replaced with common names. Since the genotypes are maintained in Wales, the

names were selected from the list of names mentioned in the Welsh literature. The list of genotypes is reported in Table C. 1.

3 Biomass characterization for biochemical conversion traits

3.1 Introduction

To replace the present oil-based energy and chemical production with plant-based alternatives, the solution must be technically possible and economically profitable. Although research has succeeded in finding technical approaches for producing fuel and platform chemicals from several sources of biomass, the application of the processes in scale is hindered by the limited profit for companies compared with the oil equivalent (Deng and Parajuli, 2016). The profitability of the conversion process of biomass can be improved by reducing the energetic and chemical inputs needed. On the biorefinery side, this means the optimization of conversion processes. In contrast, on the feedstock production side, this means producing feedstocks with suitable characteristics for higher conversion efficiencies (Li et al., 2016; Torres et al., 2019; Bhatia et al., 2020).

The efficiency of the conversion processes of the biomass into energy and chemicals mainly depends on the composition of its most abundant component, the cell wall (Li et al., 2016; da Costa et al., 2017; Da Costa et al., 2019). As much as the biochemical conversion is regarded, the quality of biomass mostly relates to overcoming the natural recalcitrance of the cell wall to deconstruction. Therefore, the most valuable quality traits of ideal biomass for the biochemical conversion is in not only in the total amount of monosaccharides contained in the cell wall but also in the strength of their interlinkages

and in the portion that can be economically released with enzymatic or physical pre-treatments (van der Weijde et al., 2013; van der Weijde et al., 2017b).

In the last decade, numerous studies have shown that cell wall composition in the genus *Miscanthus* is variable between species and genotypes of the same species (Le Ngoc Huyen et al., 2010; Lygin et al., 2011; Domon et al., 2013; Hodkinson et al., 2013; Xue et al., 2013; De Souza et al., 2015; Kalinina et al., 2017; van der Weijde et al., 2017b; Schäfer et al., 2019; Stavridou et al., 2019). The quantification of the extent of this variation is the key to the effective breeding of *Miscanthus* plants for bioenergy to tailor specific varieties to different end uses (Clifton-Brown et al., 2019). Although the genus *Miscanthus* includes 14 accepted species (The Plant List, 2013), detailed studies of monosaccharide content and saccharification efficiency, in Europe, have focused on M. x giganteus and its parental species *M. sacchariflorus* and *M. sinensis*. (Da Costa et al., 2019; Schäfer et al., 2019). The only exception is a recent study, including a genotype of *Miscanthus* floridulus (see Stavridou et al., 2019). More recently, intra-specific hybrids have been created at IBERS in collaboration with private companies, to improve the production of biomass suitable for energy and chemical production (Clifton-Brown et al., 2019). Although these hybrids have been tested for uniformity and amount of biomass achievable, they have not been tested for the biomass characteristics concerning specific conversion processes nor compared with natural accessions in terms of stress resilience, yet.

Therefore, much remains to be investigated about the extent of inter and intraspecific variation of cell wall composition and bioconversion traits relevant to chemicals and fuel production, and how they can be used for the creation of new hybrid combinations (Kalinina et al., 2017).

This chapter explores the variability in cell wall composition and quality for biochemical conversion of the 49 genotypes selected in the previous chapter. The main objectives were assessing the existence and the extent of variation related to genotype and taxon and test the relation between the composition of the cell wall in terms of monosaccharides and ability to deconstruct it enzymatically.

3.2 Material and methods

3.2.1 Plant material

3.2.1.1 Biomass sample preparation

Biomass samples of the three biological replicates of the 49 genotypes selected in 2.3.6 were provided from the *Miscanthus* group at IBERS. Samples originated from completely senesced plants harvested in early Spring 2016, and they had been processed as reported by (Huang et al., 2018). From a biological point of view, biomass from senesced plants is characterised by a higher percentage of dry matter, higher content of lignin, lower content of carbohydrate and lower ethanol yield (da Costa et al., 2017). The complete set of samples included 147 samples. At the time of sampling, plants were 4 years old. Briefly, whole plants were harvested at the height of approx. 5 cm from the soil surface and chopped to 10 cm lengths (Figure 3.1). The cut biomass (comprised of stem and leaf material) was collected in a plastic bag and weighed to determine the total fresh weight. A sub-sample of approx. 200 g was removed and its relative moisture content determined after drying to constant weight. For subsequent compositional analysis, dried samples were ground to a 1.5 mm mesh and stored in 50 mL conical centrifuge tubes (Falcon tubes) at room temperature. For the biomass characterization shown below

approximately 1 g of this biomass for each selected genotype was transferred to a new sterile 50ml Falcon tube.

3.2.1.2 Grinding

Particle size reduction is an essential initial step for proper analysis of the cell wall material, as a more coarse material can retain part of the material of the biomass that is difficult to remove with subsequent washings (Pettolino et al., 2012). Consequently, biomass material was ground with a robotic ball milling system similar to what described by (Santoro et al., 2010). Approximately 70mg were weighted in 2ml microtubes with screw cap and grip for robotic handling (Sarstedt, Cat. 72.609.001) along with two stainless steel balls. Tubes were labelled according to the origin of the identity of the material contained, placed into racks and positioned in a Plant Grinding and Preparation System from Labman® similar to the one described by (Santoro et al., 2010)available at IBERS.

3.2.1.3 Cell wall purification

3.2.1.3.1 *Alcohol insoluble residue extraction*

Cell wall material (CWM) was extracted according to (da Costa et al., 2014) with minor modifications. Efficacy of the extraction at several steps was tested, as suggested by (Fry, 1988). Cell wall purification was performed in two main phases, the alcohol insoluble residue (AIR) preparation followed by starch removal to eliminate starch interference with the following analyses. De-starching was performed with α -amylase from porcine pancreas according to (da Costa et al., 2014). The use of mammalian amylases instead of microbial amylases reduces the risk of contamination with other enzymes that can hydrolyse carbohydrates. De-starching of AIR started with the addition of 1.5mL of 70%

v/v ethanol followed by thorough vortexing (2400rpm x 30sec) and incubation for 16h in shaking incubator set at 40°C and 150rpm. Samples were then centrifuged at 10,000rpm to pellet the residue. The supernatant was aspirated, and its absorbance at 280nm (A280) was measured and recorded. Washings with 70% v/v ethanol were repeated until the reading of A280 had a stable value. Subsequently, the pellet was washed twice by adding 1.5 mL of a chloroform/methanol (1:1) solution, vortexed to resuspend the pellet (2400rpm/min x 30sec) and incubated for 30 minutes at 25°C and 150pm. After each wash samples were centrifuged at 10,000rpm for 10 minutes and the supernatant was aspired by pipette and discarded. The pellet was washed three times with 1.5mL of 100% acetone with resuspension by vortexing, incubation, centrifugation and discarding of supernatant performed as in the previous step. The samples were then dried under airflow at room temperature. After this step, dry samples were stored at room temperature until further processing.

3.2.1.3.2 *Detection of starch*

A randomly selected group of samples were tested for the presence of starch before and after de-starching with a Lugol's test, according to (Barnes and Anderson, 2017). To perform the analysis, 2mg of AIR were weighed, and 0.5 mL of Lugol's iodine solution was added. Samples were incubated for 5 minutes at room temperature, centrifuged at 10,000 x g for 10 min to pellet residue, and the supernatant was removed without disturbing the pellet. The pellet was washed with 1mL of deionised (DI) water, centrifuged discarded as previously described, and the colour of AIR was observed. Colouring was found to different extents in all the samples before the de-starching. Since no colour was seen in the AIR after de-starching, the samples were considered free from starch.

3.2.1.3.3 *Starch removal*

Starch was removed with *α*-amylase from porcine pancreas (Novozyme E-PANAA-3G) following the method from (Foster et al., 2010) as modified by (Costa, 2015). Although the previous protocols performed starch removal at 25°C, in this case, the temperature of 35°C was used considering the kinetics of the enzyme (Anitha Gopal and Muralikrishna, 2009). Starch removal was initiated by re-suspending the AIR pellet in 1.5mL of 0.1M sodium acetate buffer pH 5. Tubes were capped and incubated at 80°C for 20min. This initial step has the aim of gelatinizing the starch. Afterwards, samples were cooled on ice, centrifuged at 10,000rpm and supernatant was discarded. The resulting pellet was washed three times with 1.5mL of DI H2O, with centrifugation and supernatant removal performed after each round to remove the buffer. Next, 1.5mL of a mixture containing 0.0003mL of 1% w/v sodium azide, 0.0156mL of porcine *α*-amylase (3000U/mL) and 1.4841mL of DI H2O was added to each sample. Samples were incubated in a shaking incubator for 48h at 35°C and 110rpm. The digestion was terminated by heating the samples to 95°C for 15min; samples were subsequently cooled in ice. Finally, samples were centrifuged (10,000x g for 10 min) and the supernatant discarded. Pellet was washed three times with DI water and two times with 100% acetone as previously described and dried at room temperature under a gentle flow of air. After the drying was completed, it was necessary to break up the material in the tube with a spatula. A Lugol's staining test performed as described in the next section confirmed the absence of starch residues in the biomass material.

3.2.1.4 Total monosaccharide content

3.2.1.4.1 *Acid hydrolysis*

Total monosaccharide composition of cell wall material prepared as described in section 3.2.2 was determined by double hydrolysis (Saeman, 1945; Sluiter et al., 2012; Petit et al., 2019). Approximately 10 mg of the previously prepared CWM (Section 2.2) was weighed into 10mL Pyrex glass tubes fitted with polypropylene caps. Sample weight was recorded. Subsequently, 0.100 mL of 72% (w/w) H_2SO_4 was added, the tubes were capped and placed on a heating block set at 30°C for 60± 5min, during which time the samples were mixed using a vortex mixer every 5 to 10 minutes. A set of sugar recovery standards (SRS) was prepared and taken through the remaining hydrolysis to correct for losses due to the destruction of sugars during dilute acid hydrolysis. Subsequently, 2.777 mL of deionised water was added to dilute the acid in solution to 4% (w/w) H_2SO_4 3samples were mixed

---

3 Calculation for the amount of water to add

Density of 72%(w/w) H_2SO_4 = 1.6338 g/mL

Density of H_2O = 1 g/mL

Density of 4% w/w H_2SO_4 = 1.025 g/mL

1. The weight of 100µL of 72% w/w H_2SO_4 is:

0.1mL 72% w/w/ H_2SO_4 x d72% H_2SO_4 = 0.1mL x 1.6338 g/mL = 0.16338g 72% w/w H_2SO_4

2. The composition of 0.1 mL 72% H_2SO_4 is

0.16338 g 72% w/w H_2SO_4 x 72% w/w = 0.1176336 g of acid

0.16338 g 72% w/w H_2SO_4 x 28% = 0.0457464 g of water

to eliminate phase separation. The sealed tubes were placed in an autoclave at 121°C for 1h. Once at room temperature, the tubes were centrifuged to produce a particulate-free supernatant, and the samples were diluted ten-fold (1:10) by taking 0.100 mL of each sample and mixing with 0.900 mL of DI water. Samples were stored at -20°C until analysis.

3.2.1.4.2 *Monosaccharide quantification*

Monosaccharides in the hydrolytic solution were separated and quantified by High-Performance Anion Exchange Chromatography and Pulsed Amperometry Detection (HPAEC-PAD) respectively using a Dionex ICS5000® from Fisher Scientific operated at

---

How much water do I need to add to bring the concentration from 72% w/w to 4% w/w.

4 : 100 = 0.1176336 g : x

$$x = \frac{0.1176^{3}_{3} \times 96}{4} = 2.8232^{0}_{6} \quad g\ of\ wat^{e}_{r}$$

3. Water to add is

2.8232064g - 0.0457464 = 2.77746 g = 2.77746mL

Calculation check

3. The concentration of H_2SO_4 after dilution is:

0.1176 g acid / (2.777 g H2O + 0.16338 g 72% H2SO4) = 0.1176/2.94038=0.03999 = 3.999 % H_2SO_4 (w/w)

4. The total volume of solution present after dilution is:

(0.16338 g H_2SO_4 + 2.777 g H_2O) x (4% w/w H_2SO_4)-1 = 2.94038/1.025=2.86866 mL

45°C using a CarboPac SA10 (4×250mm) column with a CarboPac SA10G (4×50mm) guard column. An eluent generator coupled to the system continuously prepared a KOH solution at 0.001M for isocratic elution at a flow rate of 1.5 mL /min for 14min. In all cases, a volume of 0.025 ml of the sample was injected into the column and detected by PAD using a gold working electrode and an Ag/AgCl reference electrode. A set of calibration standards was prepared. The calibration curve was validated between 5 μg · mL^{-1} and 40 μg · mL^{-1}. For each monosaccharide a stock solution of 0.4mg mL^{-1} was prepared by weighing 4 mg of each monosaccharide to the closest 0.1, recording the exact weight and diluting to 10 mL. The stock solution was then aliquoted in 0.5 mL portions and stored at -20°C. Working solutions of calibration standards were freshly prepared just before analysis by thawing an aliquot of each sugar and diluting 1:10 in ultra-pure DI water to create a mixture of calibration standards at 40 μg · mL. From this initial standard, the lower concentrations of the calibration curve were obtained by sequential dilutions. An independent set of calibration verification standard (CVS) carbohydrates (D-arabinose, D-glucose and D-xylose)was prepared using a batch other than that used for the preparation of the calibration standard. CVSs were made at a concentration that fell in the middle of the validated calibration curve (10μg mL). The CVS were analyzed after each calibration set and at regular intervals throughout the sequence. The CVSs was used to verify the quality and stability of the calibration curve(s) throughout the run. Immediately before the HPAEC-PAD analysis, samples were diluted by a factor of 1:200. They had their pH increased to 7 – 9 by mixing 0.080 mL of the 1:10-diluted samples with 0.320 mL of a solution of 0.020M KOH (Costa, 2015). Samples were never allowed to exceed a pH of 9, as this would have resulted in a loss of sugars. Neutralised samples were

sometimes stored frozen up to two days before being processed. Aliquots of 0.400 mL of the diluted samples were then filtered through 0.45 µm nylon filter-vials (Thomson SINGLE STEP; Thomson Instrument Company, Oceanside, California, USA). Chromeleon software (v. 7.1; Dionex) was used for data processing. External calibration standards were used to identify and quantify the five most prominent monosaccharides detected in the chromatograms: arabinose (Ara), galactose (Gal), glucose(Glc), xylose (Xyl) and mannose (Man). These monosaccharides are the most represented in *Miscanthus* cell wall (da Costa et al., 2017) and have an impact on the cell wall recalcitrance (da Costa et al., 2019).

3.2.2 Enzymatic glucose release

Enzymatic release of monosaccharides was performed according to (Resch et al., 2015) as modified by (da Costa, 2015; Petit et al., 2019). Approximately 10mg of cell wall material prepared as per 3.2.2 were manually weighed out in 2mL polypropylene microtubes with screw cap. Exact sample weight was recorded. To each sample, 0.3 mL of 100% acetone was added to collect the material at the bottom of the tube. Acetone was let to evaporate under a stream of air overnight. To each tube, 1 mL of saccharification mixture was added having the composition shown in table 3.1.

Table 3-1 Composition of the saccharification mix in a total volume of 1 mL

Component	Amount	Final concentration
0.025M Potassium acetate buffer (pH5.6)	0.957 mL	0.0239 M
Cellulase from *Trichoderma reesei* (Cellulase, Sigma Aldrich, code C2730)	0.0024 mL	0.24%
of β glucosidase from *Aspergillus niger* (Novozyme 188, Novozyme, discontinued)	0.0006 mL	0.06%
1% sodium azide [4]	0.040mL	0.04%
Total	1 mL	

The mixture was prepared in a single batch and kept to 4°C until use. Samples were incubated for 48h at 50°C in a shaking incubator set at 150rpm. After 48h samples were taken out of the incubator, centrifuged at 10,000rpm and 0.9mL of the supernatant were transferred to a new 2mL polypropylene tube and stored for 2-3 days at -20°C until monosaccharide quantification. Just before analysis samples were diluted 1 to 50 in ultrapure DI water. Some samples were tested to assess which monosaccharides had been released by the enzyme. The solution contained quantifiable amounts of Glucose, Xylose

[4] to repress bacterial growth

and Arabinose. These monosaccharides were quantified using the Dionex system as described in subsection 3.2.1.4.2.

3.2.3 Leaf to stem ratio

To test if differences between genotypes could be a consequence of a different leaf to stem ratio between genotypes, in 2017 new plant material was collected from ABR33 to quantify leaf and stem dry weight and determine leaf to stem ratio.

3.2.3.1 Plant material

In January 2017 biomass material for the 49 genotypes selected in the previous chapter was collected from ABR33. From each plant, three full stems were collected. The material was separated in leaf and stem and then chopped in smaller pieces. Subsamples were collected in plastic bags before being freeze-dried. Weight of the whole fresh material and subsamples before and after freeze-drying was recorded for both portions. The difference in weight between the fresh and dried weight of the subsamples was used to determine the moisture of the whole plant, similar to what was done by (Huang et al., 2018).

3.2.3.2 Leaf to stem ratio determination

The calculated dry weight of leaf and stem biomass was used for the calculation of the leaf stem ratio for each genotype according to methods used in literature (Smart et al., 1998; Zhang et al., 2014a). Briefly the entire sample was separated into leaf blade sheath and stem fractions. The stem fraction potentially included stems and inflorescence (Figure 3.1).

Figure 3.1 Example of sample preparation
Three to four stems were collected from each plant (A). The material was separated into leaf and stem and portioned to fit in plastic bags before freeze drying (B).

3.2.4 Fibre composition data from Near-infrared spectroscopy

For each genotype, the content of cellulose, hemicellulose and lignin predicted by Near-infrared spectroscopy (NIR) was obtained from Dr. Sue Lister (IBERS, Aberystwyth University) from the analytical chemistry department.

Data for the biomass harvested for each genotype in 2016 was obtained from the MScan database (see chapter 2).

Data analysis

Data were statistically analysed using the R programming language (R Core Team, 2018). Along with the basic statistics available in the R environment, additional packages were required for more articulate analyses. A complete list of the packages and functions used along with the description of their use is shown in. Details of the code used, and the datasets can be found at (Link to GitHub missed). Results of the statistical analysis were

visualized following (Tuftea 1942; Tufte, 2001) and using the plot package (Wickham, 2016).

Table 3.2 List of packages used in the data analysis

Package	Reference	Function	Application
stats	(R Core Team, 2018)	Tukey HSD	Post-hoc pairwise comparison test to determine the significance
		aov	Type I analysis of variance
		lm	Linear models examination. The argument method is set to “Pearson”.
dplyr	(Wicklam et al., 2019)	summarise	Summary of the data and table creation
ggplot2	(Wickham, 2016)	ggplot	Plot creation
gridExtra	(Auguie, 2017)	grid.arrange	Multiple plot arrangement on the same page

3.3 Results

3.3.1 Selection of variables for cell wall characterization

The general aim of the experiment was to find genotypes with contrasting characteristics of the biomass in the group of 49 selected in the previous chapter. Here, the biomass of the selected 49 *Miscanthus* genotypes was characterised in terms of conversion quality and carbohydrate composition. A series of traits were measured in the purified cell wall material and stored in 19 variates presented in Table 3-2.

Variables represented four groups of characteristics of the biomass: saccharification efficiency, cell wall composition, cell wall architecture and plant architecture. The saccharification efficiency was estimated using the number of monosaccharides released upon an enzymatic treatment as an indicator. As far as the composition is concerned, it was possible to quantify the content of the 5 most abundant monosaccharides reported in the cell wall, arabinose, galactose, glucose, xylose and mannose (da Costa 2017). To represent the cell wall architecture, the ratio between cellulose and hemicellulose, xylose and arabinose, xylose and glucose and mannose and galactose were used as indicators of the arrangement of the monosaccharides in the cell wall (Held et al., 2014). Finally, the architecture of the plant was studied in terms of the ratio between leaf and stem in the aboveground biomass.

Table 3-2 List of traits measured on the cell wall material.
List and description of the variables describing compositional and bioconversion characters analyzed within the framework of this study.

	Trait	Unit	Description
Composition	Ara	mg/100 mg CW	Arabinose content in the cell wall material from the plant biomass from the, determined as D-arabinose by HPAEC
	Gal	mg/100 mg CW	Galactose content in the cell wall material from the plant biomass, determined as D-galactose by HPAEC
	Glc	mg/100 mg CW	Glucose content in the cell wall material from the plant biomass, determined as D-glucose by HPAEC
	Xyl	mg/100 mg CW	Xylose content in the cell wall material from the plant biomass, determined as D-xylose by HPAEC
	Man	mg/100 mg CW	Mannose content in the cell wall material from the plant biomass, determined as D-mannose by HPAEC
	NDF	mg/100 mg DM	Amount of Neutral Detergent Fiber predicted using NIR. Calibration was carried using the method from (Van Soest, 1963).
	ADF	mg/100 mg DM	Amount of Acid Detergent Fiber predicted using NIR. Calibration was carried using the method from (Van Soest, 1963).
	ADL	mg/100 mg DM	Amount of Acid Detergent Lignin predicted using NIR. Calibration was carried using the method from (Van Soest, 1963).

	Trait	Unit	Description
	Cellulose	mg/100 mg DM	Amount of cellulose, calculated according to the formula: $Cellu^{lo} = ADF - AD^{L}$
	Hemicellulose	mg/100 mg DM	Amount of hemicellulose calculated according to the formula: $Hemicel^{lu} = ND^{F} - AD^{F}$
	K_Lignin	mg/100 mg DM	Amount of lignin predicted from NIR. Calibration was carried using the method by
	DM	mg/100 mg	Dry matter content, determined as described in (Huang et al., 2018).
Cell Wall Architecture	AXS	Xyl/Ara %	Degree of arabinoxylans substitution expressed as the ratio between Xylose and Arabinose.
	XG	Xyl/Glc %	Glucose to xylose ratio
	GMS	Man/Gal %	Degree of galactomannans substitution expressed as the ratio between mannose and galactose.
	CelHem	Cel /Hem %	The ratio between cellulose and hemicellulose
Saccharification	AraE	mg/100 mg Ara	Amount of arabinose released by a certain amount of cell wall upon a 48 digestion with an enzymatic mix containing cellulase and β-glucosidase, expressed as a percentage of D-arabinose over Ara.

Trait	Unit	Description
GlcE	mg/100 mg Glc	Amount of glucose released by a certain amount of cell wall upon a 48 digestion with an enzymatic mix containing cellulase and β-glucosidase expressed as a percentage of D-glucose over Glc.
XylE	mg/100 mg Xyl	Amount of xylose released by a certain amount of cell wall upon a 48 digestion with an enzymatic mix containing cellulase and β-glucosidase expressed as a percentage of D-xylose over Xyl.
Ara_Enz_p	mg/100 mg CW	Amount of arabinose released by a certain amount of cell wall upon a 48 digestion with an enzymatic mix containing cellulase and β-glucosidase, expressed as a percentage of cell wall material digested.
Glc_Enz_p	mg/100 mg CW	Amount of glucose released by a certain amount of cell wall upon a 48 digestion with an enzymatic mix containing cellulase and β-glucosidase, expressed as a percentage of cell wall material digested.
Xyl_Enz_p	mg/100 mg CW	Amount of xylose released by a certain amount of cell wall upon a 48 digestion with an enzymatic mix containing cellulase and β-glucosidase, expressed as a percentage of cell wall material digested.

	Trait	Unit	Description
Plant Architecture	DWL	g	The dry weight of leaf material
	DWS	g	The dry weight of stem material
	LSR	% (g/ 100 g of DW)	Leaf to stem ratio expressed as the percentage of the leaf on the stem.

3.3.2 Saccharification efficiency

Here, the saccharification efficiency was quantified as the amount of the monosaccharides released by the CWM upon enzymatic digestion on the set of 49 genotypes. It was tested with enzymatic digestion performed with a mixture of cellulase and β-glucosidase for 48 hours at 50°C (da Costa et al., 2014). The amount of the monosaccharides released was calculated:

As a proportion (w/w %) of the cell wall and represented by the variables Ara_Enz_p, Glc_Enz_p and Xyl_enz_p. These represent the amount of monosaccharide released by the enzyme per unit of weight of cell wall, without taking in account the total content of monosaccharide.

As a percentage of the total amount of monosaccharide present on the cell wall(AraE, GlcE, XylE). This provides a measure of the efficacy of enzymatic hydrolysis not affected by the initial amount of substrate available for the reaction. This variable is more informative of the arrangement and strength of the bonds between monosaccharide units in the carbohydrates.

The test was performed with three technical replicates on the three biological replicates of each genotype, and the value was obtained by averaging the values of three technical replicates. Values were expressed as mean ± standard deviation of the mean between the three biological replicates (x ± sd) (Table D. 1). The average AraE, GlcE and XylE in the population was 5.54±3.34 %, 7.28±4.52 % and 2.22±1.112 % of the total amount of monosaccharides present and 0.191±0.262 %, 2.62±1.225%, 0.601±0.223 of the cell wall. The high ratio between standard deviation and average indicates a high variability of the three traits in the population. Therefore, the summary statistics of the variates in the single genetic groups were calculated (Appendix 2 – Table 2). The result show that, in terms of the genetic group, the highest release of all the three monosaccharides was obtained from *M. floridulus* (AraE = 6.88 ± 2.70 %, GlcE =14.61 ± 5.95 % and XylE = 3.18 ± 1.25%). The lowest release was obtained from the *M. sinensis* (EMI/PRI) (2.53 ± 0.96 %) for AraE, *M. lutarioriparius* (5.08 ± 1.69%) for GlcE and *M. sinensis* (EMI/PRI) (1.12 ± 0.29 %) for XylE. Finally, in terms of genotype the lowest of AraE was obtained from Bors (1.94 ± 0.64 %) and the highest from Angharad (10.49 ± 6.71 %), the lowest GlcE from Lamorak (2.18 ± 0.27%) and the highest from Llyr (18.65 ± 6.41%), the lowest XylE from Lamorak (0.78 ± 0.16 %) and the highest from Llyr (4.09 ± 0.52%).

In summary, these results show the high variability of the amount of the monosaccharides that is possible to release by enzymatic digestion of untreated cell wall material from different genotypes and genetic groups of *Miscanthus*.

Table 3-3 Amount of monosaccharides enzymatically released from the cell wall material

Mean and standard deviation of the monosaccharide enzymatically released from the cell wall material (CWM). The amount was expressed as a percentage in weight of the CWM for the three main monosaccharides released upon enzymatic digestion, namely arabinose (Ara_enz_p), glucose(Glc_Enz_p) and xylose (Xyl_Enz_p). Values for each biological replicate were obtained from 4 technical replicates of the experiment and averaged. Values in the table are the mean and standard deviation (SD) between 3 biological replicates (n = 3).

Genetic group	Genotype	Ara_Enz_p		Glc_Enz_p		Xyl_Enz_p	
		mean	sd	mean	sd	mean	sd
Flor_Taiwan	Gawain	0.087	0.012	3.206	0.86	0.439	0.127
	Geraint	0.242	0.019	4.547	1.118	0.753	0.113
	Lancelot	0.176	0.051	4.813	0.4	0.741	0.235
	Llyr	0.219	0.036	5.304	0.111	0.826	0.092
	Percival	0.267	0.039	5.111	1.587	0.84	0.121
	Total	0.198	0.071	4.596	1.121	0.72	0.195
Hyb.gig	Cilydd	0.076	0.014	1.703	0.347	0.424	0.059
	Lamorak	0.068	0.039	0.937	0.163	0.218	0.033
	Olwen	0.154	0.005	3.573	0.636	0.461	0.034
	Total	0.099	0.046	2.071	1.232	0.368	0.12
Hyb.mx	Galahd	0.372	0.48	2.832	0.998	0.488	0.125
	Tristan	0.183	0.053	3.93	0.698	0.782	0.17
	Total	0.278	0.323	3.381	0.977	0.635	0.209
Hyb.sin	Cigfa	0.337	0.444	2.212	0.642	0.397	0.018
	Hafgan	0.115	0.063	2.524	1.568	0.453	0.183
	Nimue	0.06	0.019	2.141	0.619	0.291	0.108
	Total	0.171	0.258	2.292	0.919	0.38	0.128

Genetic group	Genotype	Ara_Enz_p		Glc_Enz_p		Xyl_Enz_p	
		mean	sd	mean	sd	mean	sd
Lut	Blodeuwedd	0.49	0.632	1.609	0.446	0.598	0.171
	Bran	0.096	0.015	2.331	0.305	0.687	0.046
	Eigyr	0.063	0.021	2.501	0.714	0.711	0.081
	Goronwy	0.103	0.013	1.687	0.396	0.738	0.164
	Govannon	0.342	0.459	1.39	0.378	0.492	0.089
	Gwalchmei	0.103	0.029	2.525	0.575	0.942	0.33
	Gwydion	0.085	0.017	2.171	0.72	0.576	0.135
	Kay	0.392	0.451	2.791	0.426	1.132	0.155
	Lleu	0.049	0.004	1.218	0.092	0.448	0.043
	Mabon	0.099	0.008	1.64	0.352	0.626	0.127
	Math	0.117	0.008	2.326	0.401	0.792	0.088
	Owain	0.097	0.025	2.892	0.38	0.833	0.045
	Total	0.194	0.318	2.053	0.664	0.706	0.221
Sacc_rob	Aranrhod	0.521	0.729	2.109	0.52	0.692	0.198
	Beli	0.068	0.011	1.25	0.33	0.557	0.061
	Culhwch	0.083	0.016	2.58	0.398	0.728	0.151
	Gaheris	0.054	0.011	2.349	0.642	0.446	0.05
	Gwenhwyfar	0.085	0.019	2.093	0.388	0.769	0.194
	Lunete	0.104	0.011	1.898	0.608	0.701	0.212
	Modred	0.081	0.019	1.783	0.259	0.68	0.07
	Nudd	0.126	0.027	2.948	0.82	0.741	0.157

Genetic group	Genotype	Ara_Enz_p		Glc_Enz_p		Xyl_Enz_p	
		mean	sd	mean	sd	mean	sd
	Pryderi	0.467	0.563	3.187	1.979	0.854	0.214
	Pwyll	0.154	0.054	2.721	1.361	0.713	0.096
	Rhiannon	0.276	0.269	3.434	0.771	0.805	0.09
	Total	0.187	0.293	2.431	0.958	0.703	0.165
Sin (EMI/PRI)	Luned	0.071	0.021	1.984	0.271	0.376	0.085
	Total	0.071	0.021	1.984	0.271	0.376	0.085
Sin_N_Japan	Aeron	0.154	0.072	3.852	1.396	0.538	0.16
	Dylan	0.059	0.011	1.927	0.552	0.451	0.083
	Enid	0.101	0.013	2.945	0.279	0.345	0.113
	Gareth	0.236	0.023	4.968	0.67	0.664	0.189
	Peredur	0.163	0.137	1.976	0.613	0.491	0.103
	Total	0.142	0.086	3.134	1.374	0.498	0.158
Sin_S_Japan	Angharad	0.329	0.156	3.558	1.9	0.544	0.259
	Bedivere	0.377	0.409	1.77	0.161	0.274	0.083
	Urien	0.154	0.054	3.016	0.695	0.482	0.187
	Total	0.287	0.243	2.781	1.289	0.433	0.206
Sin_Taiwan	Bors	0.115	NA	1.15	NA	0.339	NA
	Bors	0.047	0.002	1.412	0.095	0.343	0.096
	Uther	0.251	0.314	2.029	0.559	0.483	0.129
	Yspaddaden	0.48	0.685	2.609	0.42	0.457	0.058
	Total	0.26	0.421	2.017	0.627	0.428	0.107

Genetic group	Genotype	Ara_Enz_p		Glc_Enz_p		Xyl_Enz_p	
		mean	sd	mean	sd	mean	sd
		0.191	0.262	2.62	1.225	0.601	0.223

Table 3-4 Amount of monosaccharides enzymatically released from the cell wall material normalised by the total monosaccharide content.

Mean and standard deviation of the monosaccharide enzymatically released from the cell wall material (CWM). The amount was expressed as a percentage in weight of the total monosaccharide content for the three main monosaccharides released upon enzymatic digestion, namely arabinose (AraE), glucose (GlcE) and xylose (XylE). Values for each biological replicate were obtained from 4 technical replicates of the experiment and averaged. Values in the table are the mean and standard deviation (SD) between 3 biological replicates (n = 3).

Genetic group	Genotype	AraE		GlcE		XylE	
		mean	sd	mean	sd	mean	sd
Flor_Taiwan	Gawain	4.725	2.984	12.06	7.484	2.108	0.719
	Geraint	7.786	0.228	10.41	3.372	2.656	0.338
	Lancelot	6.195	1.308	18.271	6.394	4.01	2.065
	Llyr	8.718	0.425	18.648	6.411	4.095	0.518
	Percival	9.929	0.535	15.977	6.38	3.776	1.11
	Total	7.35	2.429	15.406	6.26	3.377	1.294
Hyb.gig	Cilydd	4.396	0.378	3.766	0.786	1.427	0.172
	Lamorak	3.346	0.735	2.177	0.269	0.784	0.157
	Olwen	7.235	0.633	8.416	1.921	1.541	0.104
	Total	4.712	1.707	4.786	2.996	1.251	0.376
Hyb.mx	Galahd	5.947	0.533	7.166	2.195	1.725	0.407
	Tristan	9.712	1.14	9.503	0.869	2.519	0.401
	Total	7.829	2.292	8.334	1.967	2.122	0.565

Genetic group	Genotype	AraE		GlcE		XylE	
		mean	sd	mean	sd	mean	sd
Hyb.sin	Cigfa	3.128	0.622	5.285	1.875	1.22	0.112
	Hafgan	4.358	2.072	6.197	4.313	1.481	0.764
	Nimue	3.03	1.255	5.03	1.6	0.942	0.376
	Total	3.565	1.413	5.504	2.54	1.215	0.488
Lut	Blodeuwedd	6.659	3.714	3.506	0.734	1.998	0.581
	Bran	4.731	2.519	5.42	0.699	2.258	0.15
	Eigyr	4.176	2.186	6.855	1.446	2.616	0.339
	Goronwy	5.412	2.473	4.116	0.855	2.253	0.38
	Govannon	3.86	1.156	3.355	0.94	1.669	0.242
	Gwalchmei	7.173	1.027	5.94	1.161	2.948	1.064
	Gwydion	4.676	1.984	5.192	1.769	1.903	0.47
	Kay	7.497	0.696	6.959	1.522	3.581	0.537
	Lleu	2.681	0.613	3.02	0.636	1.541	0.128
	Mabon	5.351	1.639	4.013	1.179	2.149	0.499
	Math	6.601	2.774	5.62	0.974	2.536	0.147
	Owain	4.177	1.13	7.014	0.845	2.537	0.188
	Total	5.266	2.404	4.963	1.69	2.307	0.671
Sacc_rob	Aranrhod	5.437	2.225	4.617	0.924	2.248	0.638
	Beli	5.231	1.476	5.323	3.791	2.821	1.521
	Culhwch	4.266	2.561	5.532	1.044	2.237	0.658
	Gaheris	3.647	0.048	5.042	1.264	1.523	0.162

Genetic group	Genotype	AraE		GlcE		XylE	
		mean	sd	mean	sd	mean	sd
	Gwenhwyfar	4.747	1.952	4.688	0.662	2.41	0.531
	Lunete	6.027	0.318	6.596	5.046	2.995	1.731
	Modred	6.278	2.576	7.49	4.543	3.705	1.616
	Nudd	5.049	2.685	9.549	2.683	3.15	1.092
	Pryderi	10.865	8.12	11.956	4.509	4.773	2.767
	Pwyll	13.128	8.096	8.212	6.737	2.924	1.272
	Rhiannon	5.455	1.889	8.144	2.178	2.53	0.364
	Total	6.415	4.391	7.066	3.721	2.848	1.38
Sin (EMI/PRI)	Luned	2.35	0.903	5.032	0.579	1.186	0.328
	Total	2.35	0.903	5.032	0.579	1.186	0.328
Sin_N_Japan	Aeron	4.751	2.442	10.108	3.699	1.924	0.624
	Dylan	2.386	0.129	6.319	2.117	1.701	0.255
	Enid	4.372	0.492	10.021	4.37	1.702	1.149
	Gareth	6.148	0.5	13.124	1.142	2.242	0.762
	Peredur	9.421	7.251	8.976	7.886	2.128	0.551
	Total	5.44	4.06	9.71	4.419	1.918	0.642
Sin_S_Japan	Angharad	10.495	4.282	11.382	2.863	2.153	0.511
	Bedivere	2.804	0.295	4.439	0.344	0.904	0.284
	Urien	4.864	2.533	7.727	1.323	1.663	0.642
	Total	6.46	4.356	7.849	3.4	1.573	0.697
Sin_Taiwan	Bors	NA	NA	NA	NA	NA	NA

Genetic group	Genotype	AraE		GlcE		XylE	
		mean	**sd**	**mean**	**sd**	**mean**	**sd**
	Bors	1.957	0.342	3.807	0.513	1.107	0.239
	Uther	2.648	0.792	5.54	1.321	1.693	0.375
	Yspaddaden	3.993	1.08	6.051	1.668	1.375	0.276
	Total	2.866	1.132	5.133	1.495	1.392	0.365
		5.54	3.34	7.28	4.525	2.225	1.112

3.3.2.1 Effects of genotype and genetic group

The differences observed in the descriptive statistics were tested for the hypothesis that they were not related to the genotype or the genetic group (Null-Hypothesis). The assumption was tested using the Analysis of Variance test (ANOVA)(Fisher, 1919). Where the ANOVA test was significative a test for the honest significant difference (HSD) was carried (Tukey, 1949). The degree of association between groups and variables, which is known as the power of the test, was assessed by using the measure of the eta square

$$\eta^2 = \frac{SS_{effe}{}^{c}}{SS_{tot}{}^{d}}$$

The ANOVA test detected a significant effect of genotypes and genetic groups on the three variables.. However, the number of observations was not sufficient to test for the effect of interaction between genetic group and genotypes.

The genotype had a statistically significant effect on all the three variables associated with saccharification efficiency, AraE (ANOVA: F = 2.144, df =48 , p = 0.001, η2= 0.548), GlcE

(ANOVA: F= 4.032 , df = 48, p <0.0001, η2= 0.671) and XylE (ANOVA: F = 3.387, df = 48, p < 0.0001, η2= 0.631).

The results from the HSD test for AraE, GlcE and XylE are presented in Figure 3-1, Figure 3-2 and Figure 3-3, respectively. The area included between second and third quartile is highlighted. Although AraE (Figure 3-1) was within the second and third quartiles of the range of distribution of the variable for all the genotypes, some had values above and below these thresholds. For instance, Pwyll that had the highest amount of AraE, two folds the median value, and Bors, Cigfa, Dylan Llau and Luned that were in the lower quartile. The same differences were observed as much as GlcE (Figure 3-2) and XylE are concerned (Figure 3-3), genotypes that are outliers in the highest range and the low range are the same observed in AraE.

The effect of genetic group on AraE, GlcE and XylE was tested using a type I one-way ANOVA test. Although, the set of data is unbalanced (different number of genotypes for each genetic group), the ANOVA result obtained using the function anova on a linear modelobject obtained with the function lm() is reliable. However, the unbalanced design of the test resulted in a lower statistical power. The result details are presented in Appendix D - Table 5. The 10 genetic groups were significantly different for AraE (ANOVA: F= 2.51, df = 9 p = 0.011, η2= 0.155), GlcE (ANOVA: F = 13.097, df = 9, p<0.0001) and XylE (ANOVA: F = 8.38, df = 9, p <0.0001) (Appendix 3 – Table 1). Results of the significance of the comparisons according to the Tukey test between genetic groups are presented in Figure 3-4, Figure 3-5, Figure 3-6.

Notably, *Miscanthus* floridulus from Taiwan had the highest release for all the three monosaccharides. *Miscanthus* lutarioriparius and *Miscanthus* sacchariflorus robustus

had values above the median for all the three monosaccharides. Between the three types of hybrids in the group, the Hyb.mx (artificial hybrids) were always above the median value, while Hyb.gig (natural hybrids) and Hyb.sin (sin x sin hybrids) were in below the median value. These results indicate that the saccharification efficiency is a trait associated with the genetic group and that interspecific artificial hybrids have lower recalcitrance to conversion.

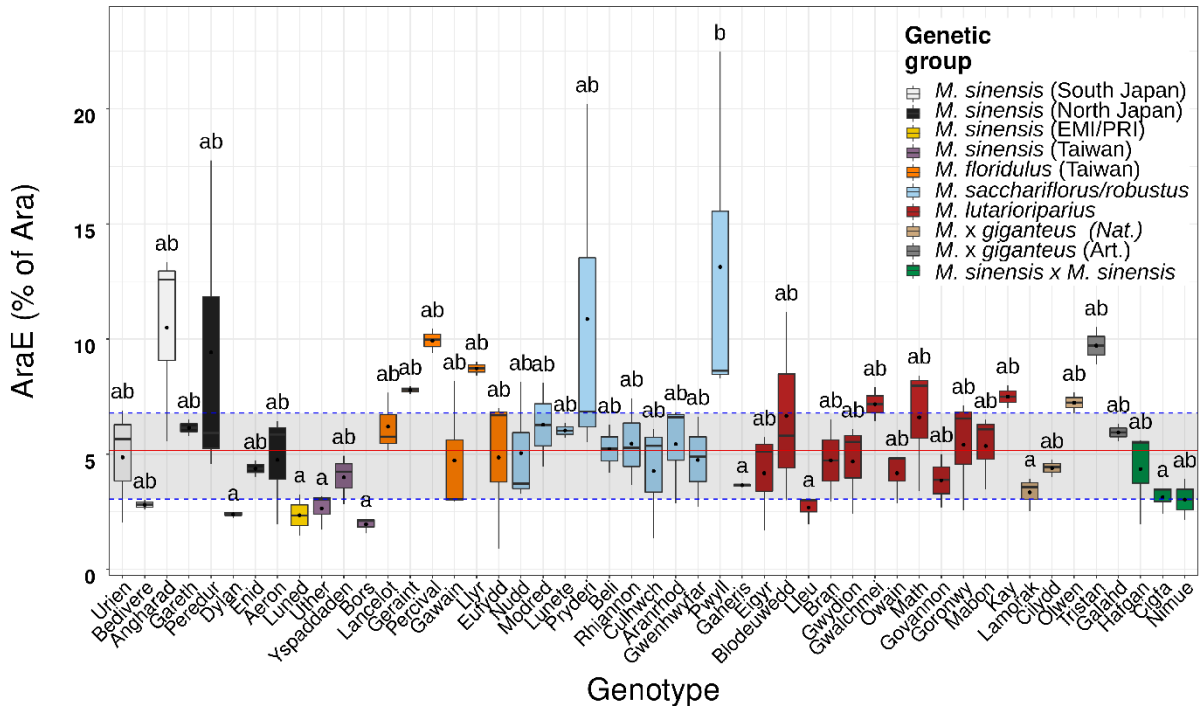


Figure 3-1 Effect of genotype on the arabinose enzymatically released from the cell wall (AraE)

Boxplot representing the effect of the genotype on the amount of arabinose released by treatment with an enzymatic mixture of cellulase and β-glucosidase for 48 hours for the 49 genotypes. Labels on the x-axes are the names of the genotypes ordered by genetic group (colour index). Values on the y-axes are the percentages of D-Arabinose enzymatically released as a percentage of the total monosaccharide in the cell wall determined by acid hydrolysis (Ara). Boxes represent the mid two quartiles with the median draw. The black dot in the boxes represents the mean values (n= 3). The horizontal black line marks the value of the median for the population. The shadowed area delimits the 2nd and 3rd quartiles of the variable in the population (n = 49). The letters represent the results of multiple pairwise comparisons using the Tukey's test (p = 0.05) for the effect of genotype on AraE. M. x giganteus (Nat.) = naturally occurring hybrids of *M. sacchariflorus* x *M. sinensis*, M. x giganteus (Art.) = hybrids of *M. sacchariflorus* x *M. sinensis* generated by crossing-operations.

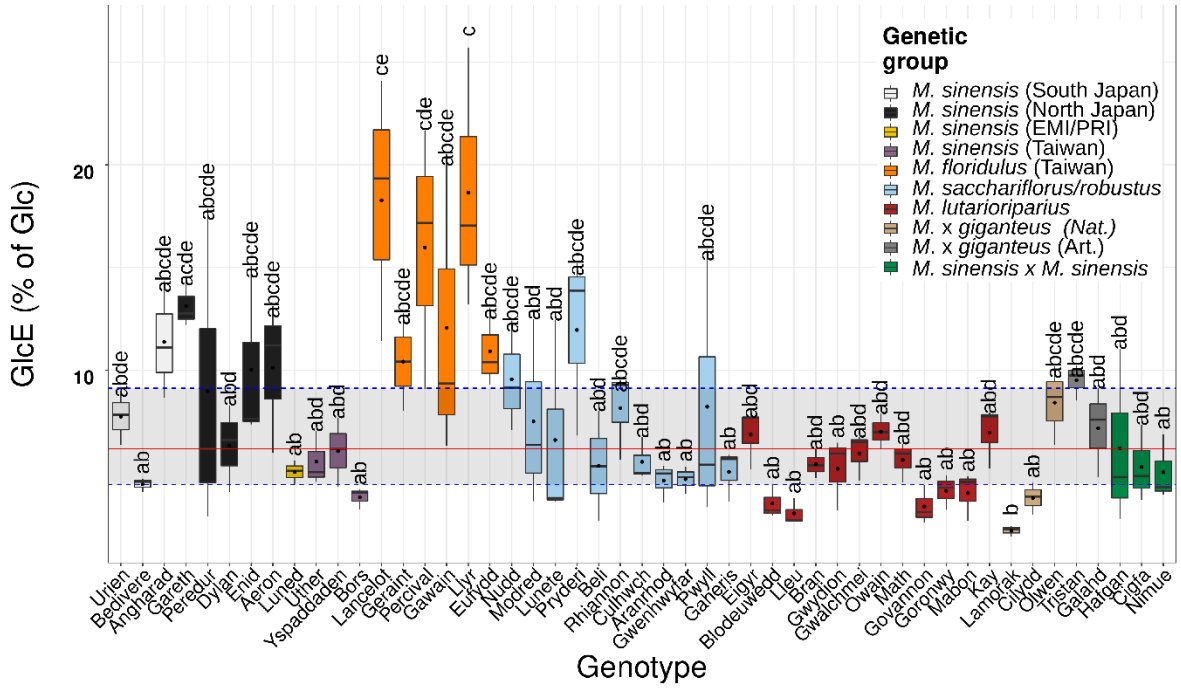


Figure 3-2 Effect of genotype on the amount of glucose enzymatically released (GlcE) from the cell wall.

Boxplot representing the effect of genotype on the amount of glucose released by treatment with an enzymatic mixture of cellulase and β-glucosidase for 48 hours for the 49 genotypes. Labels on the x-axes are the names of the genotypes ordered by genetic group (colour index). Values on the y-axes are the percentage of D-glucose enzymatically released (GlcE) as a percentage of the total monosaccharide in the cell wall determined by acid hydrolysis (Glc). Boxes represent the mid two quartiles with the median draw. The black dot in the boxes represents the mean values (n=3). The horizontal line marks the value of the median for the population (n = 49). The shadowed area delimits the 2nd and 3rd quartiles of the variable in the population. The letters represent the results of multiple pairwise comparisons using the Tukey's test (p = 0.05) for the effect of genotype on GlcE. M. x giganteus (Nat.) = naturally occurring hybrids of *M. sacchariflorus* x *M. sinensis*, M. x giganteus (Art.) = hybrids of *M. sacchariflorus* x *M. sinensis* generated by crossing operations.

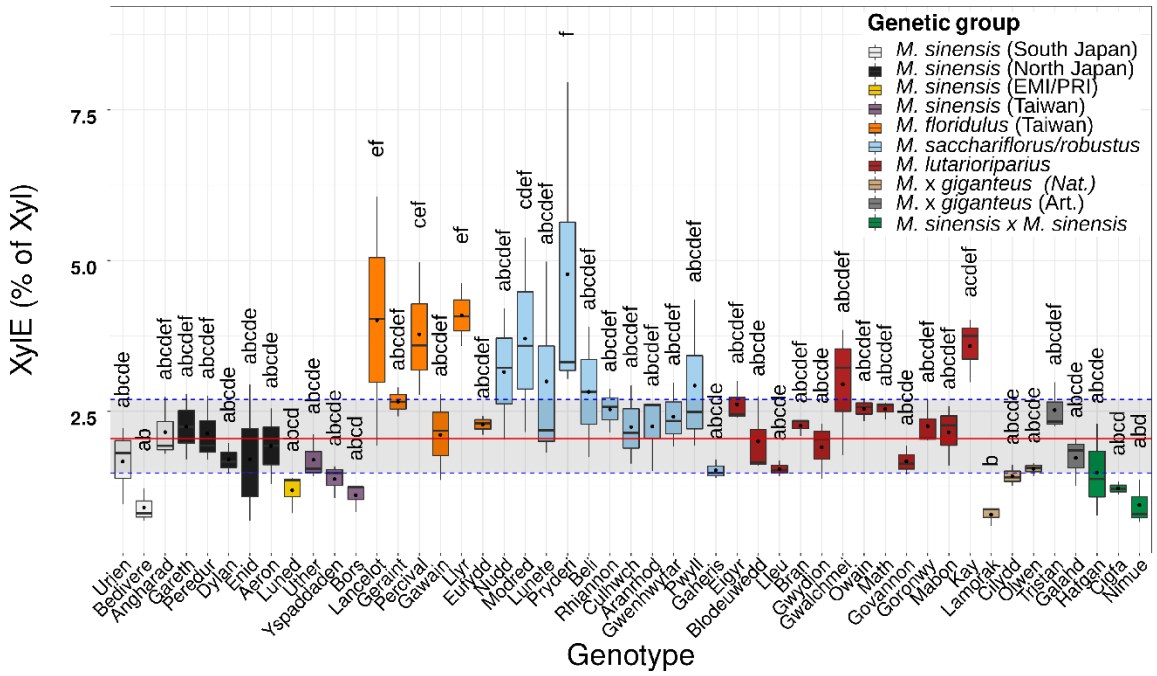


Figure 3-3 Effect of genotype on the amount of xylose enzymatically released (XylE) from the cell wall.

Boxplot representing the effect of the genotype on the amount of xylose released by treatment with an enzymatic mixture of cellulase and β-glucosidase for 48 hours for the 49 genotypes (AraE). Labels on the x-axes are the names of the genotypes ordered by genetic group (colour index). Values on the y-axes are the percentages of D-Xylose enzymatically released as a percentage of the total monosaccharide in the cell wall determined by acid hydrolysis (Xyl). Boxes represent the mid two quartiles with the median draw. The black dot in the boxes represents the mean values (n= 3). The horizontal line marks the value of the median for the population (n= 49). The shadowed area delimits the 2nd and 3rd quartiles of the variable in the population. The letters represent the results of multiple pairwise comparisons using the Tukey's test (p = 0.05) for the effect of genotype on XylE. *M.* x *giganteus* (Nat.) = naturally occurring hybrids of *M. sacchariflorus* x *M. sinensis*, M. x giganteus (Art.) = hybrids of *M. sacchariflorus* x *M. sinensis* generated by crosses.

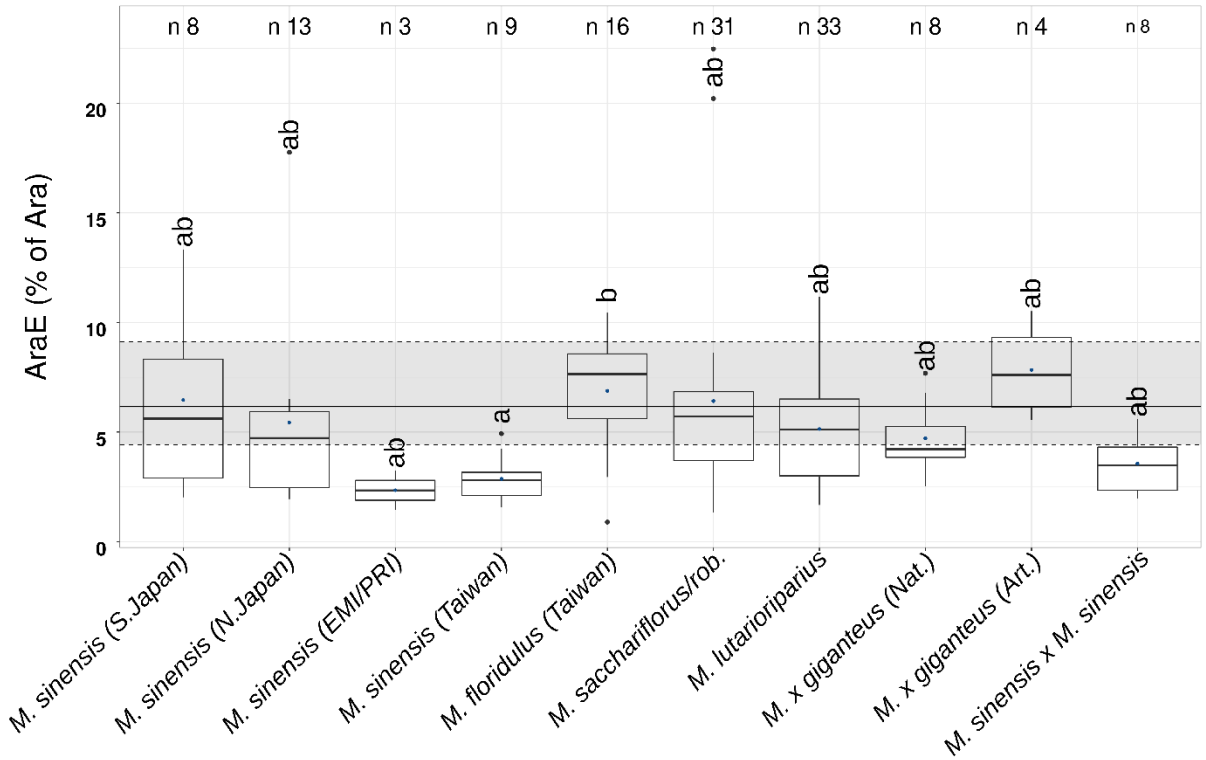


Figure 3-4 Effect of the genetic group on arabinose enzymatically released (AraE) from the cell wall

Boxplot presenting the effect of genetic group on the amount of arabinose enzymatically released by treatment with an enzymatic mixture of cellulase and β-glucosidase (AraE). Labels on the x-axes are the names of the genetic groups. Values on the y-axes are the percentages of D-arabinose enzymatically released as a percentage of the total monosaccharide in the cell wall determined by acid hydrolysis (Ara). Boxes represent the mid two quartiles with the median draw. The black dot in the boxes represents the mean values (n indicates the number of observations for each group). The horizontal line marks the value of the median for the population (n= 49). The shadowed area delimits the 2nd and 3rd quartiles of the variable in the population. The letters represent the results of multiple pairwise comparisons between groups using the Tukey's test (p = 0.05) for the effect of genotype on AraE. M. x giganteus (Nat.) = naturally occurring hybrids of *M. sacchariflorus* x *M. sinensis*, M. x giganteus (Art.) = hybrids of *M. sacchariflorus* x *M. sinensis* generated by crossing operation.

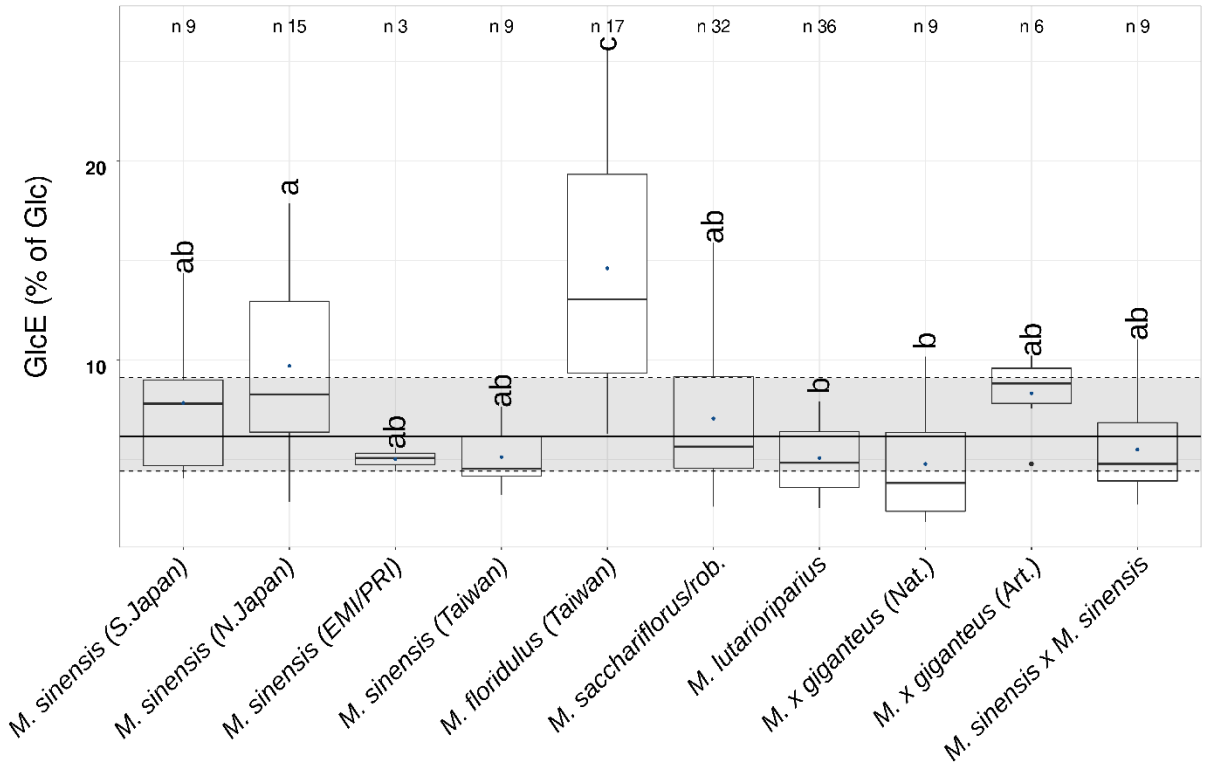


Figure 3-5 Effect of the genetic group on the glucose enzymatically released (GlcE)

Boxplot presenting the effect of genetic group on the amount of glucose enzymatically released (GlcE) by treatment with an enzymatic mixture of cellulase and β-glucosidase. Labels on the x-axes are the names of the genetic groups. Values on the y-axes are the percentages of D-glucose enzymatically released as a percentage of the total monosaccharide in the cell wall determined by acid hydrolysis (Glc). Boxes represent the mid two quartiles with the median draw. The black dot in the boxes represents the mean values (n indicates the number of observations for each group). The horizontal line marks the value of the median for the population (n= 49). The shadowed area delimits the 2nd and 3rd quartiles of the variable in the population. The letters represent the results of multiple pairwise comparisons between groups using the Tukey's test (p = 0.05) for the effect of genotype on GlcE. M. x giganteus (Nat.) = naturally occurring hybrids of *M. sacchariflorus* x *M. sinensis*, M. x giganteus (Art.) = hybrids of *M. sacchariflorus* x *M. sinensis* generated by crossing operations.

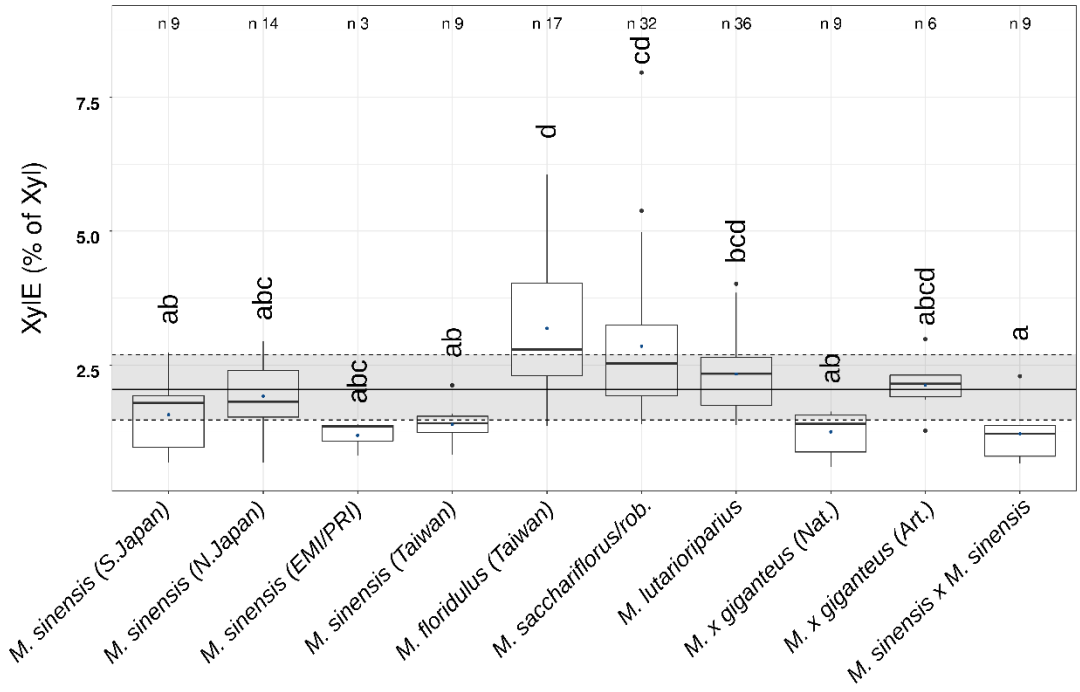


Figure 3-6 Effect the genetic group on the xylose enzymatically released (XylE)
Boxplot representing the effect of genetic group on the amount of glucose enzymatically released (GlcE) by treatment with an enzymatic mixture of cellulase and β-glucosidase. Labels on the x-axes are the names of the genetic groups. Values on the y-axes are the percentages of D-glucose enzymatically released as a percentage of the total monosaccharide in the cell wall determined by acid hydrolysis (Glc). Boxes represent the mid two quartiles with the median draw. The black dot in the boxes represents the mean values (n indicates the number of observations for each group). The horizontal line marks the value of the median for the population (n= 49). The shadowed area delimits the 2nd and 3rd quartiles of the variable in the population. The letters represent the results of multiple pairwise comparisons between groups using the Tukey's test (p = 0.05) for the effect of genotype on GlcE. M. x giganteus (Nat.) = naturally occurring hybrids of *M. sacchariflorus* x *M. sinensis*, M. x giganteus (Art.) = hybrids of *M. sacchariflorus* x *M. sinensis* generated by crossing-operations.

3.3.3 Cell Wall composition

It was hypothesised that differences in the saccharification efficiency might be related to differences in cell wall composition in terms of monosaccharide content. It was expected intra- and interspecific variation in this trait would be observed, since this is already reported in the literature for some species of the genus *Miscanthus*. A large variation between accessions of *M. sinensis* in saccharification efficiency and cell wall composition was found by Allison et al. (2011) and Zhao et al. (2014). Reviewing the information available in the literature, for *M. sacchariflorus*, *M. sinensis* and M. x giganteus, Arnoult et al. (2014) compared cell wall composition traits and reported the effect of species on cellulose, hemicellulose, lignin, and ash contents. They observed that M. ×giganteus and *M. sacchariflorus* species had more cellulose and lignin than the *M. sinensis* species, while *M. sinensis* species had higher hemicellulose content than the other two species. In addition, they observed variability in the range of variation between the species and its traits. Cellulose and hemicellulose contents had the same range of variation as for the three species, while the range of variation for the lignin and ash contents was significantly lower for the *M. sacchariflorus* species than for the M. × giganteus and *M. sinensis* species. However, there is limited information about intra- and interspecific variation assessed with experiment including all the species. *M. lutarioriparius* and *M. floridulus* have also received limited investigation regarding their cell wall traits. The data used in this study were obtained from 10 genetic groups of *Miscanthus* grown in the same experimental field location.

The cell wall monosaccharide content of each of the 49 genotypes was determined with double acid hydrolysis (Subsection 3.2.1.4) using first concentrated sulphuric acid (72%, 30°C shaking for 30 minutes) followed by treatment with diluted sulphuric acid (4%, 121°C for 45 minutes). A subsequent HPAEC analysis allowed to quantify the

amount of the main fermentable monosaccharides in the cell wall, including D-arabinose (Ara), D- Galactose (Gal), D – Glucose (Glc), D – Xylose (Xyl), D – Mannose (Man). The information was integrated with Nuclear Infrared spectroscopy (NIR) quantification data of fibre components previously performed in the chemical analysis facility at IBERS by the *Miscanthus* research group. Including this second set of data had the purpose of checking the quality of the assay and of increasing the information available to explain differences in saccharification.

The content of monosaccharides was expressed as a percentage of the amount of CWM. As expected, the most abundant monosaccharides in the cell wall were Glc (39.3 ± 7.03 %) and Xyl (28.75 ± 4.81 %). Other monosaccharides were present in lower amounts just above the limit of detection of the test, Ara (2.87 ± 3.28 %), Gal (0.8 ± 0.63 %), Man (5.52 ± 12.34 %). The ANOVA test showed that the differences in Glc and Xyl content were significant between genotypes (Table D. 6). However, the Tukey test was able to detect the differences only for Xyl. According to the Tukey test, only two genotypes (Angharad and Gaheris) were significantly different from each other, but not from the rest of the group in terms of Xyl content.

Conversely, comparing the average monosaccharide content of the genetic groups, there was a significant effect on Ara ($F = 3.713$, $df = 9$, $p = 0$, $\eta 2 = 0.212$), Glc ($F = 3.596$, $df = 9$, $p = 0$, $\eta 2 = 0.193$), Xyl ($F = 5.919$, $df = 9$, $p = 0$, $\eta 2 = 0.284$) and Man ($F = 2.082$, $df = 9$, $p = 0.035$, $\eta 2 = 0.123$). The test for comparisons using Tukey HSD allowed to localize the differences between groups (Figure 3-7, Figure 3-8, Figure 3-9). *Miscanthus* floridulus had a significantly lower amount of Ara, Glc, Xyl and Man compared with the other genetic groups. The three groups of hybrids did not differ significantly from the other groups for none of the four monosaccharides.

In summary, these results highlighted the fact that low recalcitrance occurred in genotypes with low content of monosaccharide. This result may seem unexpected. However, it is corroborated by literature reporting that linkage and arrangement of monosaccharides, more than their abundance can affect saccharification efficiency (Da Costa et al., 2019). Consequently, relations between monosaccharide, underpinning cell wall architecture were calculated.

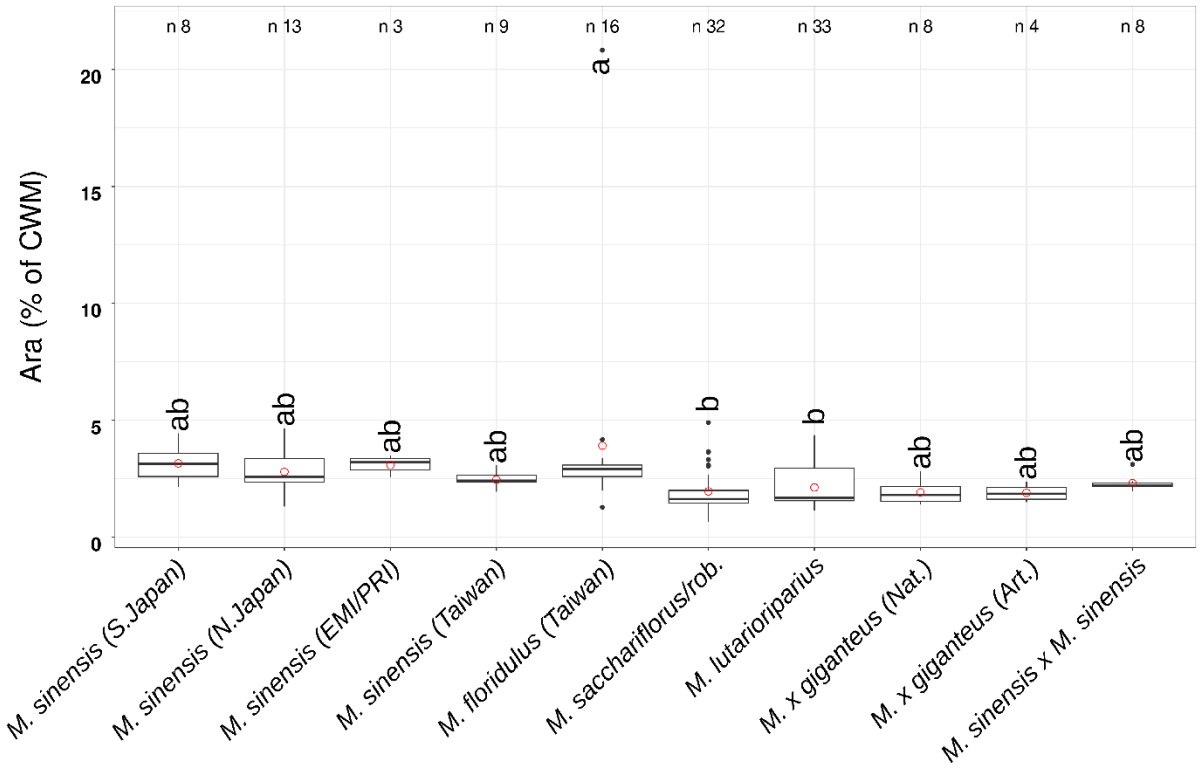


Figure 3-7 Effect of the genetic group on the cell wall arabinose content (Ara) .
Boxplot representing the effect of genetic group on the amount of arabinose released by a total acid hydrolysis (Ara) of prepared cell wall material (CWM). Labels on the x-axes are the names of the 10 genetic groups. Values on the y-axes are the percentages of D-glucose enzymatically released as a percentage of the total monosaccharide in the cell wall determined by acid hydrolysis (Glc). Boxes represent the mid two quartiles with the median draw. The black dot in the boxes represents the mean values (n indicates the number of observations for each group). The horizontal line marks the value of the median for the population (n= 49). The shadowed area delimits the 2nd and 3rd quartiles of the variable in the population. The letters represent the results of multiple pairwise comparisons between groups using the Tukey's test (p = 0.05) for the effect of genotype on GlcE. M. x giganteus (Nat.) = naturally occurring hybrids of *M. sacchariflorus* x *M. sinensis*, M. x giganteus (Art.) = hybrids of *M. sacchariflorus* x *M. sinensis* generated by crossing-operations.

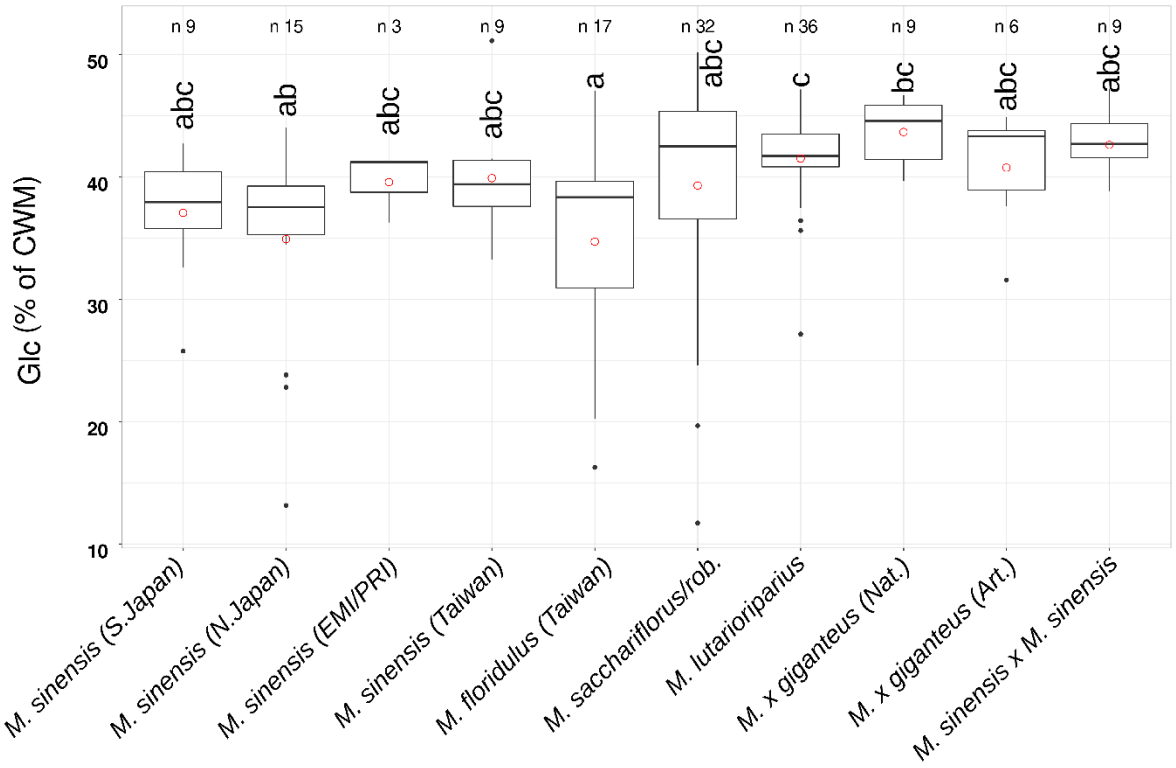


Figure 3-8 Effect of the genetic group on the glucose content (Glc) in the cell wall
Amount of D-Glucose in the CWM of the 10 genetic groups in this study. Data are means of the three technical replicates of the acid digestion. Boxes represent the mid two quartiles with the median drawn; whiskers are the 95% confidence limits plus extremes. Letters are the levels of significance according to a Tukey HSD test (p =0.05). Values on the y-axes are the amount of D-xylose as a percentage of CWM. Values on the x-axes are the names of the genetic groups (colour index).

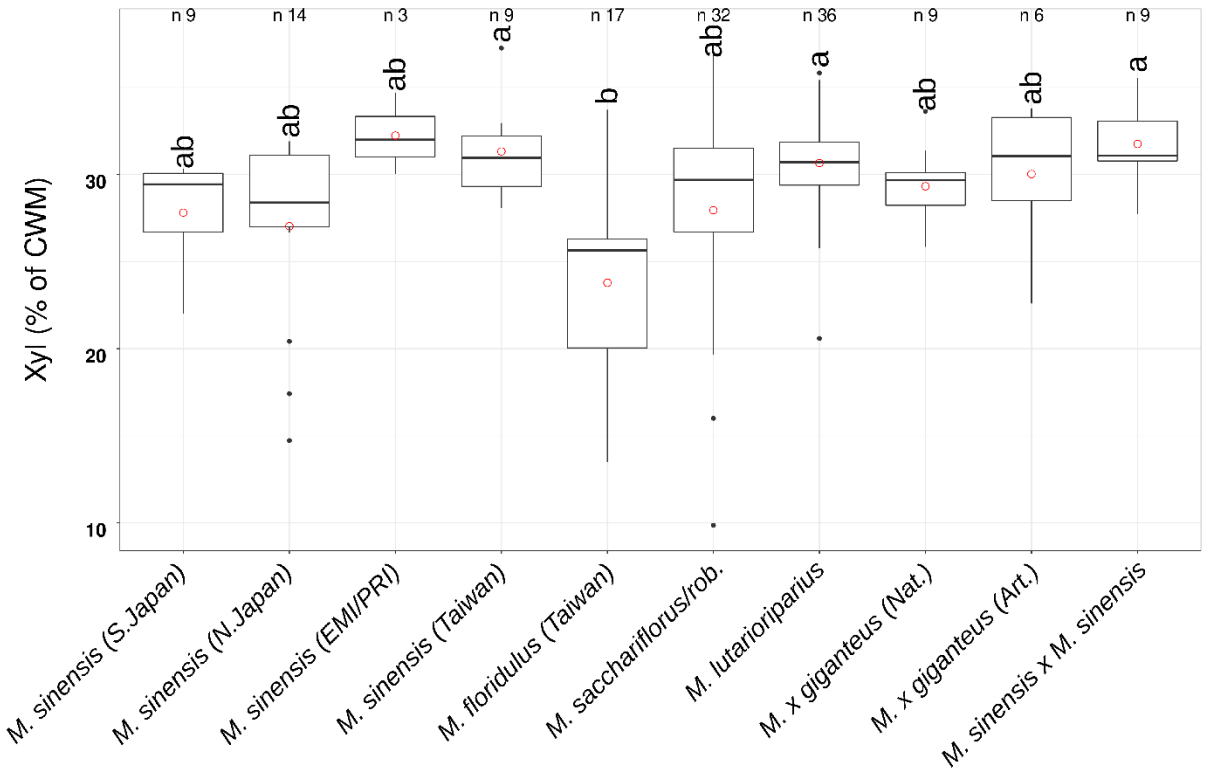


Figure 3-9 Effect of the genetic group on the amount of D-Xylose in the CWM
Xylose content in the cell wall of the 10 genetic groups in this study. Data are means of the three technical replicates of the acid digestion. Boxes represent the mid two quartiles with the median drawn; whiskers are the 95% confidence limits plus extremes. Letters are the levels of significance according to a Tukey HSD test (p =0.05). Values on the y-axes are the amount of D-xylose as a percentage of CWM. Values on the x-axes are the names of the genetic groups (colour index).

3.3.4 Cell wall architecture

The decoration of xylans in the cell wall is one of the determinants of cell wall recalcitrance (section 1.3). The ratio between the content of xylose and the content of arabinose in the cell wall can be used to describe the level of decoration of xylans in the cell wall (Wu et al., 2013; Wang et al., 2016b; da Costa et al., 2017)

Values from total monosaccharides quantification described in the previous subsection were used to calculate some derived variables describing the cell wall architecture such as the xylan decoration estimated by the proportion between xylose and arabinose (XylAra), the ratio between cellulose and hemicellulose (CelHem), the ratio between xylose and glucose (XylGlc) and mannose to galactose ratio (ManGlc). The results for the One-Way Type ANOVA test (anova on fit of linear model) for the effect of genotype and genetic group on cell wall architecture variables are presented in Table D.7. Genotypes differed significantly for CelHem ($F = 4.97$, $df = 47$, $p = 0$, $\eta 2 = 0.7$), XylAra ($F = 1.893$, $df = 47$, $p = 0$, $\eta 2 = 0.508$) and XylGlc ($F = 2.075$, $df = 47$, $p = 0.005$, $\eta 2 = 0.54$). Likewise, differences for the same variates were significant between genetic groups. The results for the posthoc test are presented in Figure 3-10, Figure 3-11. The groups of natural hybrids had significantly higher value for the ratio between cellulose and hemicellulose. The groups of genotypes of *Miscanthus* sinensis were not significantly different from each other and the hybrids. *Miscanthus* floridulus had a significantly lower value of CelHem and XylAra compared to the other groups. In contrast, *Miscanthus* lutarioriparius and *Miscanthus* sacchariflorus had significantly higher values.

Thus, these results indicate that the xylan decoration and relation between cellulose and hemicellulose in the cell wall structure of the genotypes in this study were lower in the genotypes with high saccharification efficiency (*Miscanthus* floridulus) and,

conversely, higher in the more recalcitrant genotypes (*Miscanthus* lutarioriparius and *Miscanthus* sacchariflorus).

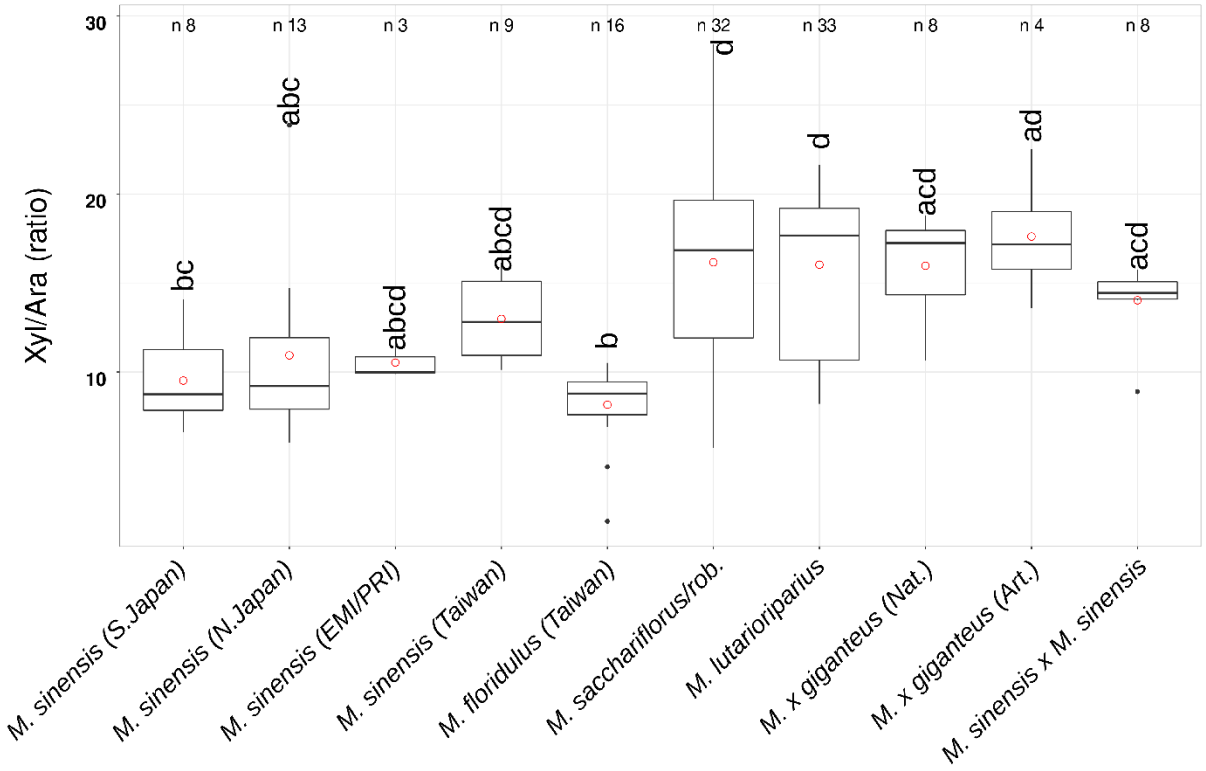


Figure 3-10 Effect of genotype on xylan decoration
Xylan decoration is represented by the ratio between D-Xylose and D-Arabinose. Data are means of the three technical replicates of the acid digestion. Boxes represent the mid two quartiles with the median drawn; whiskers are the 95% confidence limits plus extremes. Letters are the levels of significance according to a Tukey HSD test (p =0.05). Values on the y-axes are the amount of D-xylose as a percentage of CWM. Values on the x-axes are the names of the genetic groups.

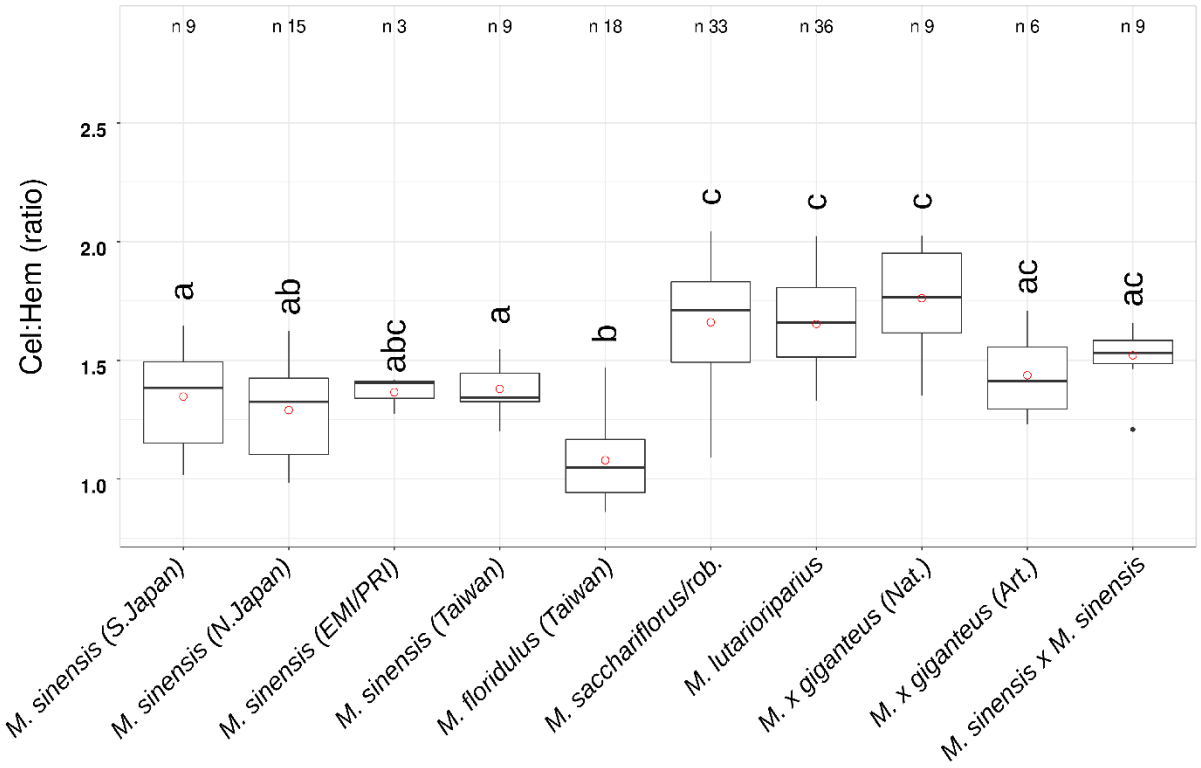


Figure 3-11 Effect of genotypes on the proportion between Cellulose and Hemicellulose.

Values used for determining the coefficient were the means of the three technical replicates. Boxes represent the mid two quartiles with the median drawn; whiskers are the 95% confidence limits plus extremes. Letters are the levels of significance according to a Tukey HSD test (p =0.05). Values on the y-axes are the amount of D-xylose as a percentage of CWM. Values on the x-axes are the names of the genetic groups (colour index).

3.3.5 Leaf to stem ratio

To test the hypothesis that differences observed between genotypes could be related to differences in plant architecture, the leaf to stem for each sample was calculated. To determine the leaf to stem ratio, new plant material was collected in 2017. The leaf to stem ratio value obtained for each genotype was the average of three stems per plant and the three biological replicates. A summary of the result is reported in Appendix 2 – Table 4. In the progression through its annual phenological cycle (Tejera et al., 2017) *Miscanthus* goes through a phase of senescence during winter. During senescence, nutrients are reallocated to the rhizome and the above-ground part of the plant dies. Senescence is accompanied by abscission of the leaves from the stem. The occurrence of the senescence phase, the timing and the percentage of leaf material lost through abscission is a trait variable between species. Some genotypes belonging to *M. floridulus* in temperate climates do not present a senescence phase. The percentage of retention of the leaves on the stem has a large range of variation between species. Here, the assessment of leaf to stem ratio was performed on material collected after Winter. Genotypes that showed a wide range of variation in the occurrence of senescence and the leaf to stem ratio of here has to be interpreted as the ability of each genotype of retaining the leaves after winter rather than a real leaf to stem ratio of the genotype. All the genotypes, except for genotypes from *Miscanthus* floridulus group, that did not senesce , had a leaf stem ratio lower than 1 (Figure 3-12). The Tukey test performed after the ANOVA showed that *M. floridulus* was significantly different from all the other groups. However, when *M. floridulus* are excluded from the comparison, it was possible to detect differences between the other genetic groups (Figure 3-13). *M. sacchariflorus* and *M. lutarioriparius* groups are not different from each other and had the lowest leaf to stem ratio. They were different from the other genetic groups.

The hybrids between sinensis genotypes and the interspecific hybrids both had a leaf to stem ratio tendentially higher than the naturally occurring hybrid.

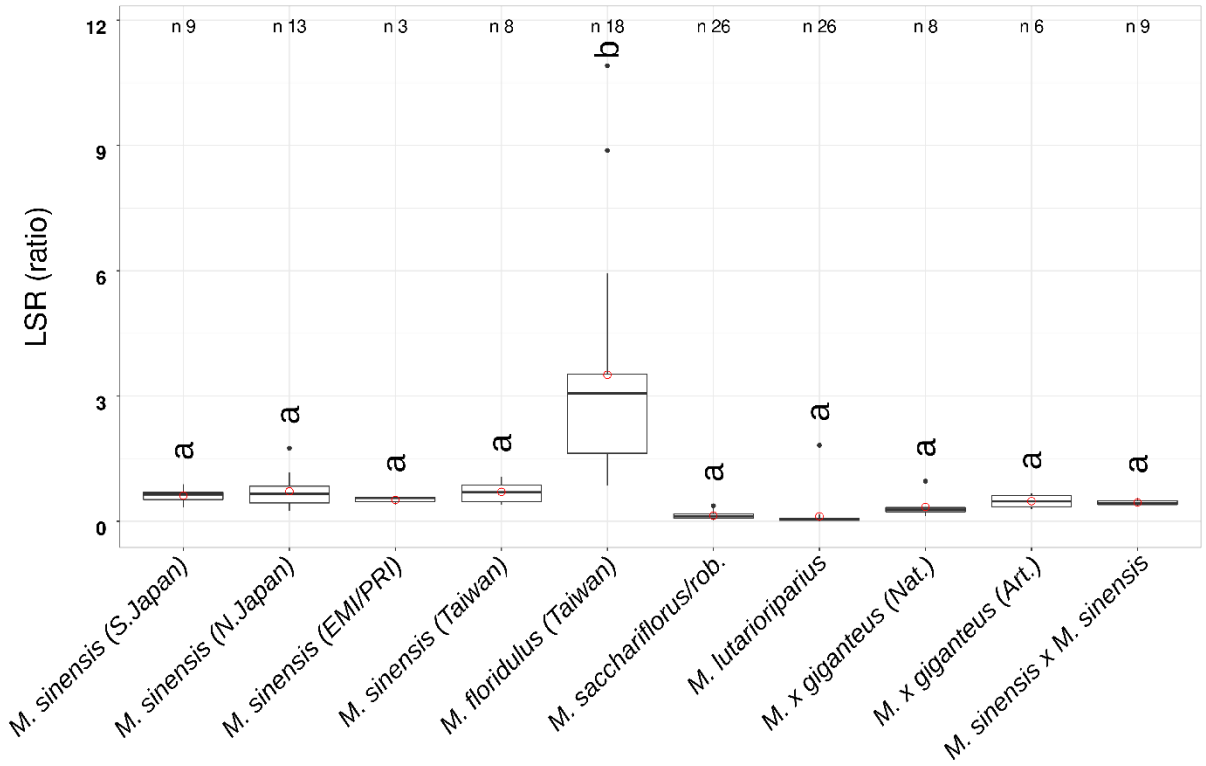


Figure 3-12 Effect of the genetic group on the leaf to stem ratio.
Boxes represent the mid two quartiles with the median drawn; whiskers are the 95% confidence limits plus extremes. Letters are the levels of significance according to a Tukey HSD test (p =0.05). Values on the y-axes are the leaf to stem ratio. Values on the x-axes are the names of the genetic groups (colour index). When all the genetic groups are considered, the effect on LSR was significant only in *M. floridulus* compared to all the other groups.

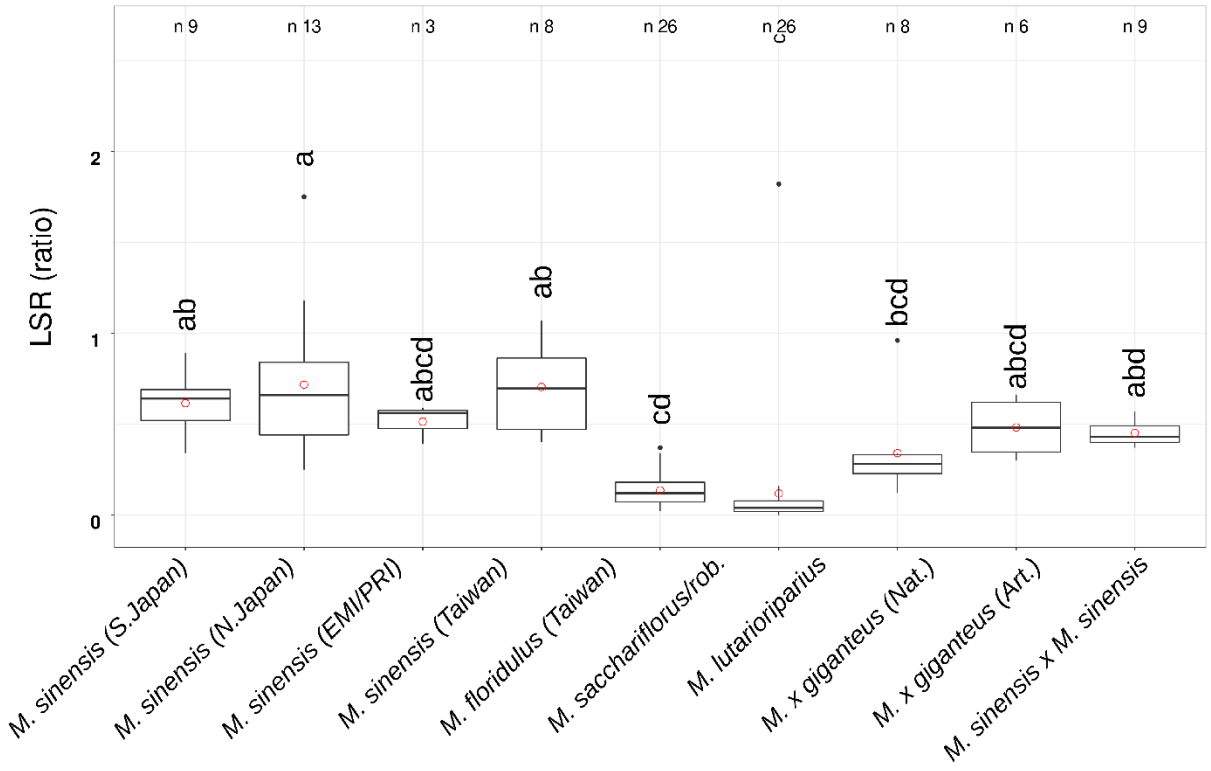


Figure 3-13 Genetic group effect on LSR lower than 1.

Boxes represent the mid two quartiles with the median drawn; whiskers are the 95% confidence limits plus extremes. Letters are the levels of significance according to a Tukey HSD test (p =0.05). Values on the y-axes are the leaf to stem ratio. Values on the x-axes are the names of the genetic groups (colour index). When only the genetic groups with an LSR lower than 1 are considered, the effect of genetic group on LSR was significant between groups.

3.3.6 Relations between cell wall composition, architecture and saccharification

The hypothesis that differences observed in the saccharification efficiency could relate to cell wall composition, cell wall and plant architecture, was tested using regression techniques between variates.

A principal component analysis (PCA) was carried out on the set of data to simplify the collection of data, identify the primary sources of variation and the relations between variates. The first two components were able to explain 48.6% of the variability present in the dataset. The plot of the PCA loadings (Figure 3-14) allows seeing the contribution of each variate to the total variance of the group, while the plot of the scores for each genotype (Figure 3-14 A) allowed to see the contribution of the individuals to the variability

Observing the loadings (Figure 3-14), it can be seen that the variates with the highest positive contribution to the first component were NDF, ADF, ADL, K_lignin, along with CelHem and XylAra. The highest negative contribution to the first component was from Hemicellulose, GlcE, and Glc_enz_p together with Ash. The variables contributing positively to the second component were Xyl, Glc and Ara, while the XylE and the AraE were contributing negatively. In brief, the first component describes variates correlated to the cellulose degradation process, and the second represents the variates involved in hemicellulose degradation. Some loadings are traversal to the quadrant suggests that a third component could be involved in explaining the rest of the variance of the dataset. Turning to the score plot (Figure 3.18)), it can be seen that the individuals spread mostly along with the first component. When marked with a colour level proportional to their GlcE, although there is no grouping, it can be seen that genotypes with low GlcE are on the right side of the score plot, while individuals with high glucose are on the left with a gradual transition from right to left. This result

agrees with the fact that the primary enzyme used was a cellulase and suggested that the variables contributing to the first component are playing a role in determining the saccharification efficiency of the genotypes.

The PCA results suggested the existence of a correlation between independent variates in the dataset. Thus, correlations were further investigated by calculating their reciprocal regression coefficient using the Pearson method for each combination of variates. Results of the test are presented in the correlation matrix in Figure 3.19.For each variate, the matrix contains a plot of its distribution and a test of regression with all the other variables. For each combination, an r-square value is reported in the figure. Only combinations where the values were higher than 0.45 were considered in an attempt to select the independent variables able to explain half of the variability in the dependent variable. . For these, the level of significance of the correlation was calculated. If we look at the plot from top to bottom, it is possible to note some interesting relations. Ara correlates with Gal (R = 0.57) and XylAra (R = -0.84). The correlation between Xyl and XylAra was instead very low. This result suggests that the amount of xylose is mainly responsible for the variation in the ratio between Ara and Xyl. Glc is correlating positively with Xyl (R= 0.87) and negatively with GlcE (R = -0.68) and XylE (-0.57). Also, Xyl relates negatively with GlcE (R = -0.64) and XylE (R = -0.59). These results together suggest that Xyl and Glc are related and variating in the same extent and direction amongst genotypes, and they are negatively affecting the ability to deconstruct the cell wall enzymatically. Also, Xyl seems to have a negative correlation with AraE (R = -0.43). Given the previous observation, it is not surprising to see that GlcE and XylE are varying together (R = 0.67). This result also points to the fact that increasing the glucose release may increase the release of other components of the cell wall and vice versa. The amount of glucose released was negatively affected

by ADF (R= -0.56), NDF (R= -0.53), ADL (R= -0.55), Klason Lignin (R= - 0.43), Cellulose (R = - 0.53), CelHem (R= -0.52), XylAra (R= -0.48) and positively by Ash. Interestingly the content of hemicellulose did not correlate with GlcE. These results are a numeric quantification of what seen in the PCA. They may indicate that the organization of monosaccharides in the cell wall played an important role in determining the recalcitrance to deconstruction in the genotypes. Klason lignin correlated positively with CelHem (R= 0.80). This could be seen as a tendency of lignin to decrease where ratio between cellulose to hemicellulose increases and vice versa and may indicate a possible competitive structural role between lignin and hemicellulose in their association to cellulose. Cellulose and Hemicellulose relate to each other (R= 0.62) and both to the CelHem (R = 0.90). CelHem, in turn, is also correlating positively to the XylAra. This relation may suggest a role of xylan decoration in the relation between cellulose and hemicelluloses. None of the above variables had a correlation coefficient with LSR higher than 0.40 apart DWL. These results suggest that while the variability in the amount of leaf material is determining the differences observed in LSR, the effect of LSR on the quality of pooled biomass is limited. All the relationships listed above had a p-value lower than 0.05, and their details are reported in respective regression plots in Appendix D.

In Figure 3-16, variates were scaled between genotypes and used to determine the distance between genotypes and variates (hierarchical clustering). The grouping was obtained using the hclust() function in R. The default method of clustering used by the function is the farthest neighbour clustering. It defines the distance between two clusters as the maximum distance between the single components. In the beginning, each individual is a separate cluster. Successively, they are aggregated to larger clusters until they are all included in a single group. The technique allowed to

investigate relations between genetic groups in terms of biomass composition. Variates were aggregating in five distinct clusters (Figure 3-16 Clustering of cell wall variates in the 10 *Miscanthus* genetic groups.

The correlation was based on the cell wall traits from the 49 genotypes in this study. The calculation was carried with the R function r-clust(). Colours represent the intensity of each variate in each genetic group. Values of variates were normalised across genotypes., top dendrogram):

- NDF and DM
- LSR, CelHem, XylGlc, Gal, XylE, Ara and Ash
- Man, ADL, AraE, GlcE
- ManGal, DWL, XylAra, Klason lignin
- DWS, Xyl, Hem, ADF, Glc, Cell

The groups are in agreement with the correlation seen above for the variates.

genetic groups were aggregating in 5 groups:

- *Miscanthus* sinensis from Japan, together to artificial hybrids, and *Miscanthus* sacchariflorus.
- *Miscanthus* sinensis EMI/PRI coming from Utukushiga, Japan, together with *Miscanthus* sinensis from Taiwan, China.
- Natural hybrids together with intraspecific hybrids
- *Miscanthus* floridulus
- *Miscanthus* lutarioriparius

These groupings indicate that, in terms of biomass composition, some observations are worth to be mentioned. Artificial hybrids group together with the parental genotypes. Natural hybrids group far from them and closer to intraspecific hybrids. *Miscanthus* sinensis EMI/PRI is closer to *M. sinensis* from Taiwan rather than

Japanese accessions as expected. *Miscanthus* lutarioriparius has biomass characteristics distinct from the other groups. In summary, the clustering observed was in agreement with genetic and geographical patterns known in the dataset.

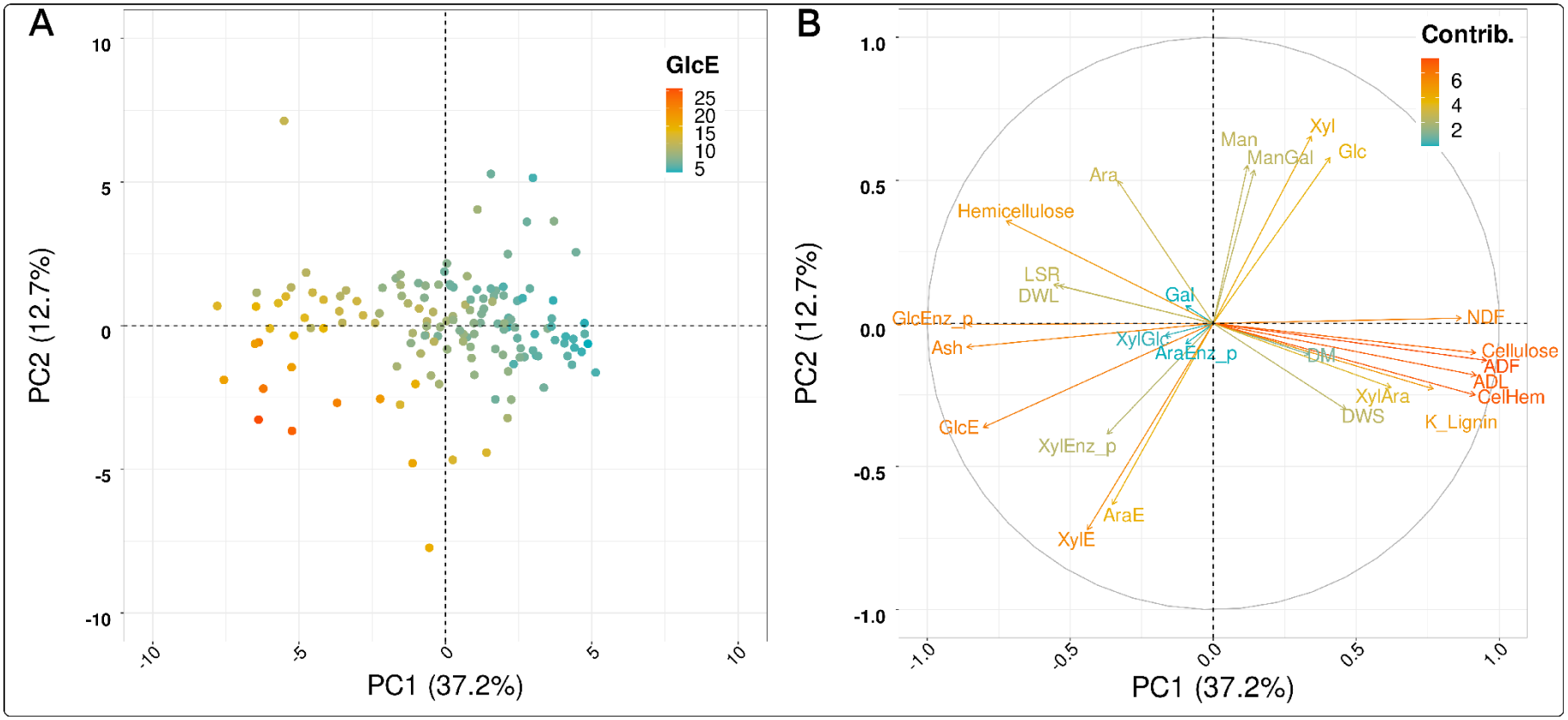


Figure 3-14 Principal component analysis (PCA) of cell wall variables.
The principal component analysis was performed on the variate set of all the traits from the 49 genotypes of *Miscanthus* spp. The arrows represent the contribution of each variate to the total variability of the dataset. Length and colour represent the contribution of each variate. The angle between the arrow line and the component axes is proportional to their correlation.

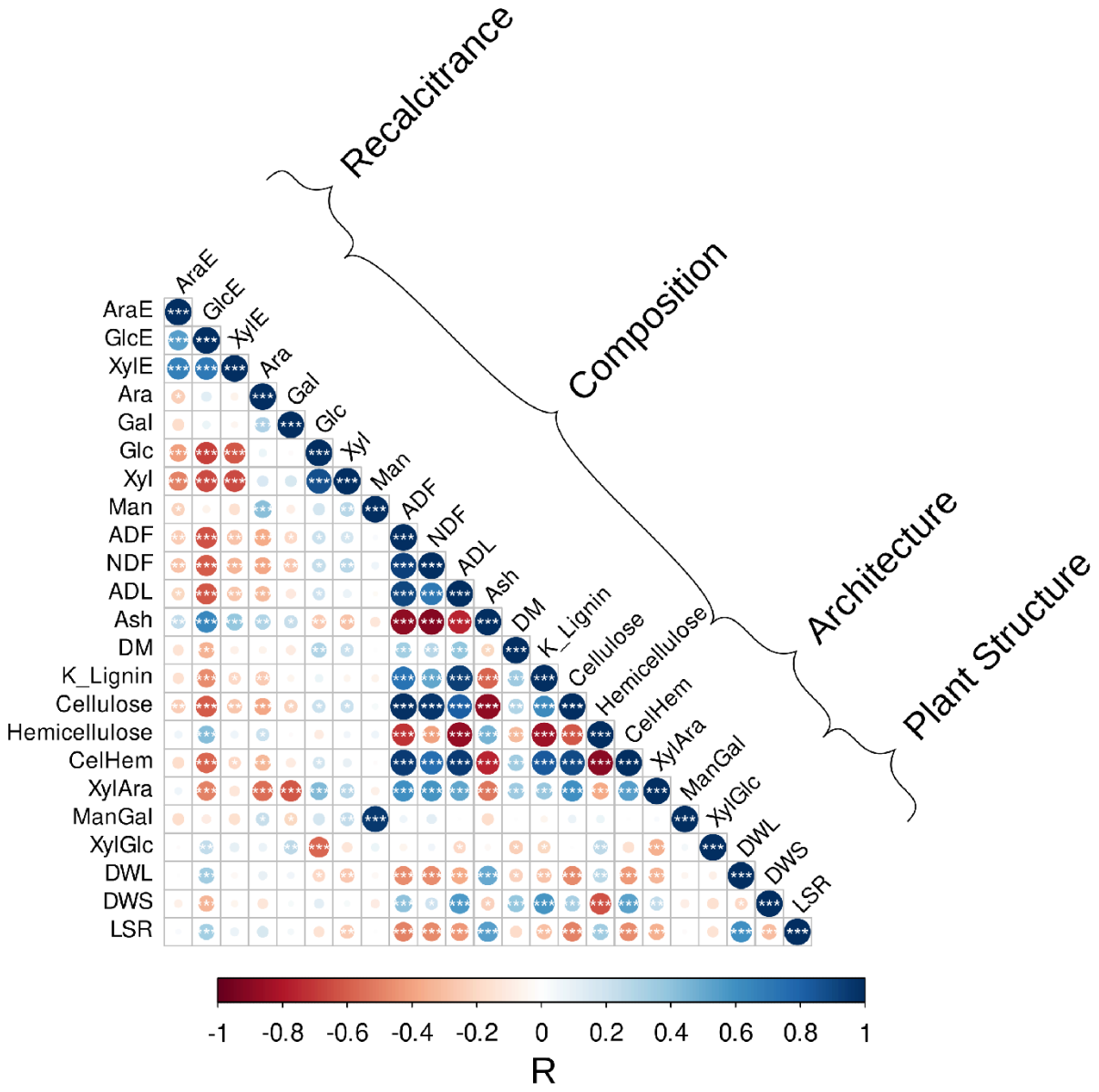


Figure 3-15 Correlation between the set of cell wall variables measured.
The correlation between the variates was tested using a Pearson correlation method with Bonferroni correction. Sign.: "***" .001, "**" .01, "*" .05, " " = n.s.

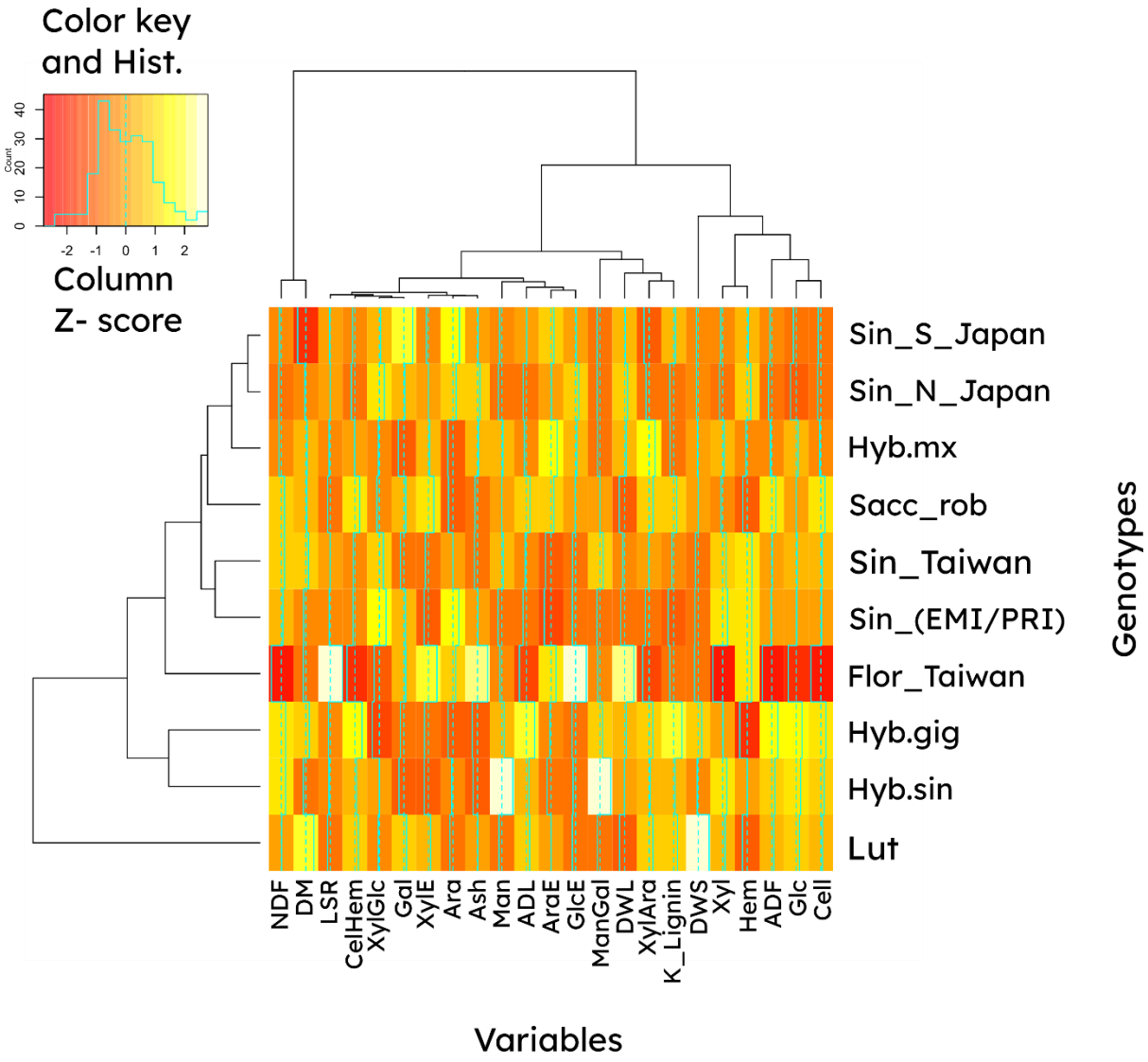


Figure 3-16 Clustering of cell wall variates in the 10 *Miscanthus* genetic groups. The correlation was based on the cell wall traits from the 49 genotypes in this study. The calculation was carried with the R function r-clust(). Colours represent the intensity of each variate in each genetic group. Values of variates were normalised across genotypes.

3.4 Discussion

3.4.1 Saccharification efficiency

Previous studies have investigated the differences in the ability to release monosaccharides with enzymatic treatment between genotypes in the genus *Miscanthus* (van der Weijde et al., 2017b; Adams et al., 2018; Da Costa et al., 2019). Although they focused on the hybrid *M. x giganteus* and its parental *M. sacchariflorus* and *M. sinensis* genotypes. In the present work, the efficiency of enzymatic conversion of cell wall material into fermentable monosaccharides was evaluated for 49 genotypes

of *Miscanthus* belonging to 10 genetic groups. There were three groups of hybrids distinct in naturally occurring hybrids (Hyb.gig), artificially interspecific hybrids (Hyb.mx) and artificial intraspecific hybrids (Hyb.sin).

Results indicated that the recalcitrance to the deconstruction of the cell wall of the 49 genotypes was mostly associated with the genetic group to which they belonged. The effect the genetic group on saccharification efficiency in *Miscanthus* is not new and was shown before by other studies. For example, Cerazy-Waliszewska et al., (2019) investigated the bioethanol production of M.x giganteus, *M. sacchariflorus* and *M. sinensis*, and found significant differences between genetic groups, but also between genotypes of the same group. In particular, in this case, *M. sinensis* and M.x giganteus were the best performing and producing almost twice as much of the amount of ethanol of the other groups. The same differences between the same groups in saccharification efficiency are found in the present study. However, here the difference is put in perspective compared with a higher number of genetic groups. The same can be said when comparing the present study to the work of Adams et al. (2018) and da Costa et al. (2014).

Also, in the present study, a significant difference in saccharification efficiency was found between groups of hybrids. Artificial hybrids (Hyb.mx) in this study had lower recalcitrance to conversion compared to the intraspecific and natural hybrids groups. This is the first report of differences in saccharification related to the type of genetic combination considered.

Genotypes of *M. floridulus* from Taiwan had the highest amount of glucose and xylose enzymatically released. This result was in agreement with the previous literature. In a study aimed to compare the saccharification efficiency of different *Miscanthus* species, Liu et al., (2013) showed that *M. floridulus* and *M. sinensis* had the highest rates of

sugar released. However, the study did not include a comparison with any hybrid accession. Up to our knowledge, there is no report of direct comparisons between *M. floridulus* and *Miscanthus* hybrid accessions. The genus *Miscanthus* includes several species, the production of interspecific crosses is hindered only by the flowering time for many combinations and the existence of more natural hybrids cannot be excluded. In summary, further investigation could help to fully exploit the natural variability available in nature and develop a more variable range of hybrids for specific conversion needs.

3.4.2 Total monosaccharide content

In the attempt to better understand the differences observed between genotypes in their ability to release monosaccharides, the composition of their cell wall in terms of monosaccharide content was investigated.

Although there was a low statistical power in the differences in cell wall composition, between genotypes, their trend mirrored the saccharification efficiency negatively. The values observed were comparable with those previously reported in the literature (Lee and Kuan, 2015). Previous studies with a lower number of genotypes, using immunological tests for sugar quantification have shown genotype related differences also in other groups of monosaccharides (De Souza et al., 2015).

Although it was not possible to detect differences between genotypes in monosaccharide content of the cell wall, there were significant differences at the genetic group level. Interestingly, *Miscanthus* floridulus group, that was the best performing in terms of monosaccharide release was the only one significantly differing from the other groups in terms of the monosaccharide composition of the cell wall. Noticeably, genotypes of *M. floridulus* in the study by (Liu et al., 2013) also had the lowest content of monosaccharides in the cell wall, which may indicate that structural

differences rather than monosaccharide content are playing a central role in their higher saccharification efficiency.

Hybrids, both natural, artificial, intra- and inter-specific, did not differ significantly from the other genotypes and between each other. The results do not seem surprising since other authors have pointed to the fact that interlinkages between cell wall fractions play a more prominent role in determining the recalcitrance of cell wall to deconstruction. Indeed, (Da Costa et al., 2019) results showed that crystallinity of cellulose instead that abundance and acetylation of AraXyl were the main traits influencing the cell wall deconstruction negatively.

3.4.3 Cell wall architecture

Cell wall architecture was investigated using some variables describing the reciprocal ratio of monosaccharides in the cell wall. However, this approach can be considered reliable only for providing an indication of cell wall architecture based on the reciprocal proportion of monosaccharides.

Results showed that the xylan decoration and relation between cellulose and hemicellulose in the cell wall structure of the genotypes in this study were differing only for *Miscanthus* lutarioriparius and *Miscanthus* sacchariflorus. This includes the fact that there was no significant difference between the groups of hybrids. In 2016 a study characterised the family of glycosyltransferases (GT) involved in the decoration of xylans in *M. lutarioriparius* (Wang et al., 2016a). Differences in the expression of GT expression between genetic groups have never been investigated. They may represent a possible target for the breeding of *Miscanthus* as a crop.

Leaf to stem ratio

The ratio between leaf and stem could be responsible for differences in the final quality of the biomass for conversion due to the different cell wall composition. In this study,

according to the leaf to stem ratio, two main groups were identified. One, including *M. floridulus*, having a ratio higher than one and considered to be leafy genotypes. The second, including all the other genotypes having a ratio lower than one.

M. sacchariflorus and *M. lutarioriparius* groups are not different from each other and both different from the other genetic group had the lowest leaf to stem ratio.

3.4.4 Relations between plant architecture, cell wall structure, cell wall composition and enzymatic cell wall deconstruction

Furthermore, the relationships between the variables representing saccharification efficiency, cell wall composition and cell wall architecture in the cell wall of the 49 genotypes were investigated using a PCA and hierarchical cluster analysis.

The PCA allowed finding two components able to explain 49.9% of the variability present in the set of variates. The first component was relating to the cellulose degradation process and the second to the hemicellulose degradation. This is particularly important, considering that the ability to degrade cellulose and hemicellulose has a primary role in the use of lignocellulosic biomass for biorefining.

According to the result of the PCA and the regression analysis, organization of monosaccharides in the cell wall played a role in determining the recalcitrance to deconstruction in the genotypes. The observed distribution of genotypes with different GlcE on the first principal component in relation with most of the variates concerning cellulose recalcitrance can be explained by the composition of the enzymatic mixture used for the digestion of the cell wall. The enzymatic mixture contained cellulose-degrading enzymes, while hemicellulose degrading enzymes were not added.

More information was derived from the regression matrix between the set of variates (Figure 3-15). GlcE was correlating negatively with ADL, Klason Lignin, NDF, ADF, cellulose, DM, Glc and Xyl and positively by AraE and XylE. *Miscanthus* biomass with high hemicelluloses levels has a lower cellulose crystallinity and higher digestibility.

In contrast, *Miscanthus* biomass with high levels of cellulose shows increased crystallinity and low biomass saccharification (Xu et al., 2012).

Another factor that negatively correlated with GlcE was the XylAra. The effect of xylan decoration on the biomass digestibility was reported for the first time in 2013 (Li et al., 2013). The authors attributed the effect of xylans decoration on saccharification efficiency to its action on the interaction between arabinose and cellulose. Interestingly, in this study genotypes belonging to *M. floridulus* were those with the lowest content of glucose and xylose as a percentage of CWM and those able to release the highest amount of glucose and xylose enzymatically at the same time.

It has been shown that low extent of arabinose decoration of arabinoxylan-oligomers (AXOS) can increase their cellulase inhibition activity (Xin et al., 2019). It has been hypothesised that AXOS have inhibitory activity by binding the active site of the enzyme. Arabinose may be hindering the ability of AXOS to bind the active site and reduce their inhibitory activity. In the same way, it is possible to think that natural differences in AXOS decoration occurring in the genotypes in this study can affect their saccharification quality. This result points to the necessity of a more detailed analysis of xyloglucan decoration differences between genotypes and genetic groups. From the correlation analysis in this study, the variation in the ratio between arabinose and xylose relates to the amount of xylose . Accordingly, cellulose and hemicellulose significantly related to each other and both to the CelHem. CelHem in turn is also correlating positively to the XylAra. This result may suggest a role of xylan decoration in the relation between cellulose and hemicelluloses. It has been shown, at least in *Arabidopsis thaliana*, that the pattern of xylan decoration is crucial for the interaction between xylans and cellulose surface (Grantham et al., 2017).

The leaf to stem ratio and the amount of XylE showed no correlation while the correlation was significant ($p<0.001$) but with low r-squared value in the case of GlcE. This indicates the role of the plant architecture in determining the quality for conversion.

Ash is known to have a negative effect on the quality of biomass for thermic conversion. Ash from biomass has a composition substantially different from coal. The use of biomass as a feedstock in thermo-reactors designed for coal generates several problems such as ash fouling and slugging. Also, metals in the biomass ash are responsible for many unwanted reactions (Demirbas, 2005). In the present study, the amount of GlcE determined by the saccharification test correlated positively with Ash determined by the NIR model. This finding is in apparent contrast with the observation that minerals contained in the ash have an inhibitory effect on cellulase activity (Bin and Hongzhang, 2010). However, it has been previously observed that silicon content, a component of plant ash, in the biomass is positively correlated with cell wall saccharification as mutants of Brachypodium distachyon with reduced silicon uptake were concentrating this mineral in the cell wall and having a lower conversion efficiency compared with the wild type (Głazowska et al., 2018). This result points to the practical difficulties of having *Miscanthus* varieties able to produce biomass for multiple conversion processes. In turn, this may hinder the ability to provide farmers with types of *Miscanthus* for various conversion processes and could represent an extra cost for the biorefineries that may transport biomass from far to optimise their operations. However, the selection could benefit from the effect of the leaf to stem ratio on conversion quality.

Valuable information was obtained from the hierarchical clustering of the genetic groups according to cell wall variates collected here. Clustering was mirroring known

genetic and geographical patterns in the dataset. Artificial hybrids grouped with the parental genotypes, while natural hybrids grouped far from them and closer to intraspecific crosses. As far as M.xgiganteus and M.sacchariflous are concerned, Da Costa et al., (2019) found the same type of clustering. However, that study was including only genotypes of *M. sinensis* and sacchariflorus. There are no other studies in literature where the same approach has been reported of a more significant number of genetic groups.

3.4.5 Selection of 12 genotypes

The main aim of the experiments described in this chapter was to identify a group of genotypes out of the original group of 49 with different characteristics in terms of saccharification efficiency and cell wall composition that could be used for further investigating effects of plant architecture and abiotic stress on the cell wall. The group included 12 genotypes. The group was selected to include genotypes from the broadest range of genetic groups, differing in saccharification efficiency and having contrasting cell wall structures using the variables measured in this chapter.

The group included three hybrids, Lamorak, Galahad and Tristan. Lamorak is a naturally occurring hybrid, previously classified as *M. lutarioriparius* (MScan, accessed on 27th of February 2020). Coming from China, it was obtained from CERES. The particular genotype used in this study comes from a progeny of 2700 seeds harvested on the 15th of January 2015 at IBERS. Conversely, Galahad and Tristan are two artificial hybrids sharing the same parents. They were obtained from the same cross operation in Venlo on the 18th of March 2011. The female parent was a *M. sacchariflorus*, and the male parent was a *M. sinensis* from North Japan. Both the parents are available in ABR33, but they were not included in the group of 49 selected. Royalties cover hybrids, and further information cannot be published. However, it can

be requested contacting the owner of the royalty directly. Considering the performance in terms of enzymatic deconstruction, the selection of 12 genotypes also included four genotypes of *M. floridulus*, Lancelot, Gawain, Geraint and Percival. The reason for the inclusion of these genotypes was their different characteristics in terms of biomass. Lancelot and Percival released 30% more glucose when treated enzymatically. However, the composition of their cell wall in terms of monosaccharides was not significantly different (Figure 3.2). The selection also included two *M. sinensis*, Gareth from North japan and Bedivere from South Japan (Figure 3.2). The first having a significantly lower release of glucose upon enzymatic treatment. Finally, a *M. sacchariflorus* and a *M. lutarioriparius* (Gaheris and Kay) with a value of GlcE close to the median value, were included in the selection.

For these genotypes, the result of the characterisation of leaf and stem characterisation is presented in chapter 4. In contrast, the test for compositional changes induced by nutrient stress is presented in Chapter 6.

Table 3-5 List of 12 genotypes selected for further experiments.

Name	Ploidy	Country	Genetic group
Lancelot	2	Taiwan	Flor_Taiwan
Gawain	2	Taiwan	Flor_Taiwan
Geraint	2	Taiwan	Flor_Taiwan
Percival	2	Taiwan	Flor_Taiwan
Bors	2	NA	Sin_Taiwan
Lamorak	2	China	Hyb.gig
Kay	2	China	Lut
Gareth	2	Japan	Sin_North_Japan
Bedivere	2	Japan	Sin_South_Japan
Gaheris	4	S Korea	Sacc_Rob
Galahad	2	NA	Hyb.mx
Tristan	2	NA	Hyb.mx

4 Biomass characterization for pyrolytic conversion

4.1 Introduction

Thermochemical conversion methods, e.g. fast-pyrolysis and gasification, offer an alternative to enzymic deconstruction and industrial fermentation for the production of fuels and chemicals from biomass. Among the thermochemical processes, this study has focused on the analysis of experimental samples under pyrolysis to investigate their behaviour in thermochemical conversion conditions. Pyrolysis is the thermic decomposition of organic matter carried at a temperature between 300 and 600°C in the absence of oxygen (Venderbosch and Prins, 2011). The technique was already in use in ancient Egypt, where the product was used for sealing boats (Fahmy et al., 2020). During pyrolysis, biomass is converted to a mixture of char, a mix of no condensable gases, and condensable vapours (liquid). When the biomass (mainly wood) is heated slowly to 300-400 °C, a slow pyrolysis takes place, and its main product is charcoal. Fast pyrolysis, instead, is realised by rapid heating to 500 °C followed by rapid cooling and condensation of the vapours produced. In this latter case, the primary product is a dark-brown mobile liquid, usually called bio-oil. The yields of bio-oil for continuously operated laboratory reactors and pilot plants are generally in the range of 60–70 wt% from a woody feedstock (Venderbosch and Prins, 2011) and around 36 wt% for grasses (Suntivarakorn et al., 2018).

The development and optimization of novel pyrolytic processes for the production of chemicals and transportation fuels have been the focus of considerable research interest in the last decade (Wang et al., 2017). The chemical composition and

properties of pyrolytic liquid, solid and gaseous products depend wholly on the biomass used as feedstock and the reaction conditions. The first report of the effect of reaction conditions on the features of the yield from pyrolysis was published in 1875 (Gruner, 1875). It is only in the last decades that the relationship between biomass composition and the products obtained by its pyrolysis has been clarified in any detail. However, the complexity of the plant cell wall hinders our ability to understand the behaviour of biomass during pyrolysis (Wang et al., 2017). What is clear is that the response of the biomass to pyrolytic conditions is related to its main components and their interactions (Pang et al., 2014). Recently, the pyrolytic behaviour of cellulose, hemicellulose and lignin has been reviewed by (Yang et al., 2007; Wang et al., 2017). Pure hemicellulose decomposes between 220 and 315℃ with the maximal rate of decomposition occurring at around 268℃.

The products obtained from the hemicellulose depend on its monosaccharide composition (Wang et al., 2017). The pyrolysis of hemicellulose from hardwood and straw results in the formation of acids, furfural and five-carbon compounds. In contrast, pyrolysis of hemicellulose from softwood produces six-carbon compounds, such as HMF and anhydrous sugars (Wang et al., 2017). The pyrolysis of pure cellulose happens at a higher temperature (315–400 °C) with peak rates of decomposition occurring at around 355 °C. The pyrolytic products obtained from cellulose are mostly furan and levoglucosan. Their yield can be influenced by the degree of polymerization of its molecules (Katō and Komorita, 1968; Mettler et al., 2012). Amorphous regions of the cellulose produce HMF and furfural between 200 and 300 ℃ (Antal, 1983). Lignin decomposition takes place at a prolonged rate, starting at relatively low temperatures and continuing up to 900℃. This slow rate of decay is mainly because lignin is thermochemically stable, having an aromatic backbone.

The products obtained from pyrolysis of lignin are influenced by the relative content of three basic units of the lignin, namely p-coumaroyl (4-hydroxycin-namyl), coniferyl (3-methoxy 4-hydroxycinnamyl) and sinapyl (3,5-dimethoxy 4-hydroxycinnamyl) alcohols, which are also known as p-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units (Figure 1.X). The mutual bonds and the linkage patterns between these three components also affect the condensation degree of lignin pyrolytic degradation products. The broader range of degradation of lignin is a consequence of the successive cleavage of linkages as the reaction temperature increases (Wang et al., 2017). Moreover, the functional groups present in the structure influence the characteristics of the products from lignin pyrolysis. For example, methoxyl groups in lignin lead to the formation of lignin char. The hydroxyl groups at the Cα or Cγ position act as hydrogen donors during pyrolysis, thus affecting the reactivity of the linkages (Kawamoto et al., 2007).

Moreover, it was shown that increased lignin content in the feed-stock led to an increased amount of total bio-oil yield, with lower water content and, in turn, a higher heating value (Fahmi et al., 2008). The ash and water content of the biomass also influence the products obtained from the pyrolysis of biomass (Venderbosch and Prins, 2011). It has been shown that as the yield of ash increases there is a concomitant reduction in the yield of bio-oil and an increase in the yield of char, which indeed retains most of the inorganic minerals that comprise the ash (Fahmi et al., 2008; Wang et al., 2017).

The pyrolytic behaviour of a feedstock can be investigated by Thermal Analysis (TA). In TA, samples of biomass are heated isothermally, or by a suitable ramped temperature program in a flow of inert gas, e.g. oxygen-free nitrogen, or argon and the weight loss is recorded. The function representing the weight change as a function of

the temperature is defined thermogravimetric analysis (TGA). The earliest reports of this technique to analyse materials is attributed to Antoine Laurent de Lavoisier (Lavoisier 1743-1794, 1790). When the weight change is compared to reference material, it is possible to use the derivative of the TGA to produce a differential thermal analysis (DTA). When the two analyses are performed simultaneously (DT-TGA), the thermal analysis is defined as simultaneous thermal analysis or STA. The DT-TGA provides numeric information (e.g., mass losses, specific temperatures) and signals shape information (shape of the curve) (Feist, 2015). DT analysis also allows the measurement of enthalpy changes during the reaction time (enthalpimetry). However, the precision of an enthalpimetric DT measurement has a limited accuracy because of the strong dependency from experimental conditions. A more precise measurement of the changes of enthalpy during thermogravimetry is achieved by using a Differential Scanning Calorimetry (DSC). DSC quantifies the change of temperature in the sample in comparison with a reference. The heat flow (HF) as a function of temperature provides useful information about structure transitions in the sample. Typical transitions in the samples include glass transition, an endothermal process, cold crystallization, an exothermic process; and melting of the crystalline parts.

The use of TG-DSC for analysis of biomass samples allows the mass fraction of the sample to be calculated, i.e. the relative proportions of moisture, volatiles, fixed carbon and ash in the sample. Such information is termed proximal analysis and is a standard requirement for feedstocks to be used for combustion or other thermolytic processes at the commercial scale. The gas evolved (EG) can be characterised using gas chromatography, perhaps coupled with mass spectrometry (GC-MS), or by Fourier Transform Infra-Red spectroscopy (FTIR), absorption spectroscopy that helps characterise the composition of the volatile, and to some degree, the pyrolytic gas

products. This approach has been extensively used for the characterisation of *Miscanthus* biomass in the last decade. TG-GC-MS was used to identify products of pyrolytic degradation of *Miscanthus* (Hodgson et al., 2011; Greenhalf et al., 2013). Hodgson et al. (2011) used Py-GC–MS to investigate the pyrolytic behaviour of *Miscanthus* biomass and found differences in the products obtained between genotypes and growing stages. The techniques were also used to compare the quality for pyrolysis of *Miscanthus* biomass with spruce wood and wheat straw (Butler et al., 2013). This study reported wood was a more desirable feedstock for pyrolysis and upgrading applications compared to *Miscanthus* or straw. They reported a high ash content in *Miscanthus* straw, resulting in the bio-oil having lower organic yield and higher water content. This also led to phase separation in bio-oil, compromising both storage and use as a feedstock for subsequent conversion.

When FTIR spectroscopy is used to analyse pyrolytic products, the results are more challenging to interpret, mainly when applied to chemically complex samples. Generally, it is not possible to identify and quantify pyrolysis products directly in such samples, in contrast to an analysis by TG-GC-MS. However, even with complex samples FTIR gives information about pyrolytic processes, and by use of chemometric models based on y data from GCMS it is possible to use FTIR data to quantify discrete components, follow gas evolution process and identify reaction intermediates of biomass pyrolysis. FTIR has been used to characterise the pyrolytic characteristics of cellulose, hemicellulose and lignin (Yang et al., 2007), celluloses with different structures (Xu et al., 2020), and various types of biomasses including tea (Tian et al., 2016).

Characterisation of *Miscanthus* biomass for thermal conversion has focused on *M. x giganteus*, *M. sacchariflorus* and *M. sinensis*. A comparison with a broader range of

other species of the *Miscanthus* genus in terms of pyrolytic characteristics has not been reported yet.

The experiment described in this chapter aimed to expand knowledge of the pyrolytic behaviour of *Miscanthus* biomass to species of this genus that have been rarely if ever, included in previous investigations of this type. Also, it was intended to test the hypothesis that TG-DSC features could be used to predict the quality of biomass for other conversion processes such as biochemical conversion. Finally, the study tested the hypothesis that TG-FTIR spectra of biomass contain information correlated to the genetic background and geographical origin of the samples. The 49 genotypes selected in Chapter 2. and characterised for biochemical conversion in Chapter 3 were used for this study.

4.2 Material and methods

4.2.1 Plant material

The plant material used in this section was described in 3.2.1.1. However, here the whole ground biomass, without cell wall extraction, was used.

4.2.2 Instruments

4.2.2.1 Thermogravimetric analysis and differential scanning calorimetry

The thermogravimetric measurements were performed using a TGA/DSC3 apparatus (Mettler Toledo, Switzerland) equipped with an autosampler able to handle 33 samples per run. The method and temperature used were adapted from Acquah et al. (2017). Around 7 mg of dry ground, biomass was placed in 70 μL aluminium oxide crucibles (Mettler Toledo, Switzerland). The crucible was weighed before and after the addition of the sample automatically by the autosampler. The weight was automatically recorded. The auto-weighting was performed in an aerobic atmosphere. As soon as the weighing was completed, the gas flow was switched to Argon gas. The

speed of the gas was maintained at 70 cm^3/min. The duration of each analysis was approximately 40min per samples (Figure 4-1). Samples were heated from 40 to 580 °C with a temperature ramp program. In detail, the temperature was held at 40 °C for 3 min and then raised to 105 °C at a rate of 25 °C /min. The temperature was then maintained at 105 °C for 5 min, after which it was increased to 580 °C at a rate of 25°C/min (see Figure 4-1).

A cleaning cycle in which no sample was present in the crucible was carried out every 15 samples. After each run, the crucibles were cleaned in a furnace. The furnace was heated to 700°C. The temperature was kept constant for 1 hour. This insured the complete cleaning of the crucibles, to avoid the effect of the residues on the next run.

4.2.2.2 Fourier Transformed Infra-Red (FTIR) spectroscopy

The composition of the gases emitted during the TG program was tested by FTIR. The FTIR of the emitted gases was performed using a Shimadzu IRTracer 100 spectrometer (Shimadzu, Japan).

The gases were transferred from the TG system to the FTIR using a heated line at 200°C.

The spectra were collected between 700 and 4000 nm from the start of the TGA (40°C) for 40 minutes at automatic regular intervals to obtain 35 FTIR spectra per run.

4.2.3 Data analysis

4.2.3.1 TGA and DSC data

TGA and DSC data were collected and exported using STARe evaluation software, available online from Mettler Toledo (https://www.mt.com/de/en/home/library/software-downloads/lab-analytical-instruments/STARe_Eval_SW.html). The exported files were imported in R environment for analysis. The files were joined in a data frame with 2576 variates

(mass loss at different time points) and 141 observations. A principal least square regression (PLSR) model was built to test if the mass loss spectra were able to predict biomass characteristics collected in chapter 3. The model was constructed in R environment using the pls package. The limited number of samples available (141) did not allow for the validation of the model on a test set of data. Therefore, the model was validated using a cross-validation (CV) with 10 randomly generated segments. The optimal number of components for the model was the one that optimised the root mean squared error of prediction (RMSP).

The quality of the model was evaluated using the R square value between observed and predicted values according to Equation 4.1.

Equation 4.1

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2}$$

4.2.3.2 FTIR data

FTIR spectra were exported in the proprietary format of Shimadzu. Subsequently, they were renamed accordingly and converted to text format using Essential FTIR software (available at http://www.essentialftir.com/index.html). The file from each sample was a 6844 x 34 matrix, whose values were the absorbances at each point. The data were imported in R environment. In total, 119 files out of 147 were validated and used for the analysis. Next, the spectra were scaled, centred. For the analysis, only spectra from time 11 to time 22 were considered. The times deemed to correspond to the maximum

weight loss in the TGA. Finally, data frames were bound together in a 6844 x 34 x 119 matrix, which was used for the multivariate data analysis.

For the analysis of the three-way, matrix generated two approaches are possible to reduce the dimensionality of the dataset. On the one hand, the array can be unwrapped to a lower degree matrix such as:

$$K \cdot I + J$$

Another approach is represented by parallel factor analysis, PARAFAC/CADENCOMP or CP. The CP is a particular type of principal component analysis where the problem of choosing a specific orientation of the axes is solved by the principle of parallel proportional profiles as a basis for determining such orientation (Cattell, 1944). The advantages of using a CP are the unicity of the solution because of the rotational degrees of freedom and the possibility of understanding the structure underlying a dataset without pre-knowledge. The number of components to use in the model was estimated using a CORCONDIA diagnostic method as described by Bro and Kiers (2003) using the function corcondia() from the multiway package in the R environment.

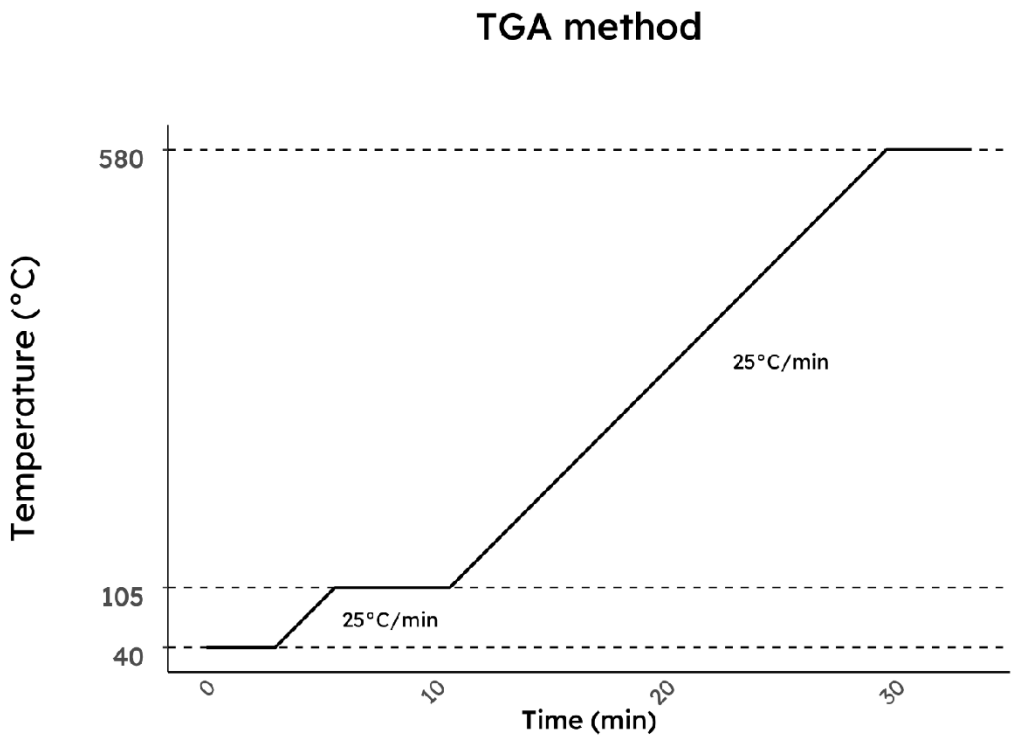


Figure 4-1 Temperature program used in the TGA
The thermal behavior of the biomass from the 49 genotypes. The reaction was happening in a non-oxidative environment created by purging the air from the system with a constant flow of Argon gas. Selection of temperature was made adapting the method used by Acquah et al. (2017)

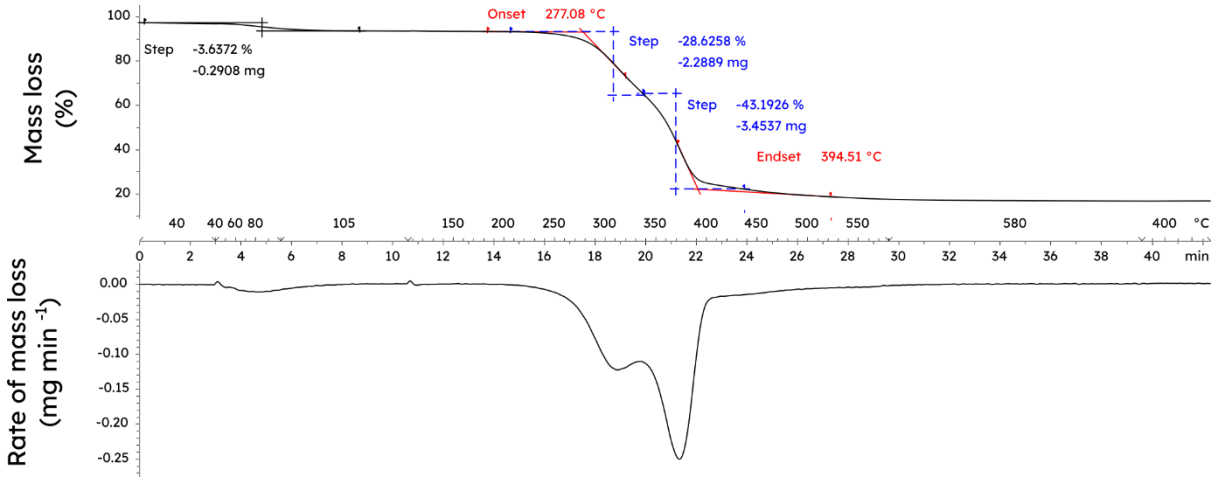


Figure 4-2 Variables obtained from thermogravimetric analysis and differential scanning calorimetry
Parameters measured are here exemplified in one spectrum. The mass loss (Top) and the rate of mass loss (Bottom; its first derivative) are plotted as a function of the time and reference temperature. The same measurements were taken on all the 294 samples obtained from two technical replicates of the TGA on the 147 samples.

Table 4-1 Variables describing each sample TG curve.

Variable	Unit	Description
step 1	wt %	Expressed as a percentage of the total sample weight. Loss of weight between 100 and 120°C due to the dehydration of the sample.
onset	°C	The temperature at which the maximum loss of weight begins.
end set	°C	The temperature at which the maximum loss of weight finishes.
step 2	wt %	Expressed as a percentage of the total sample weight. Loss of weight in the area of the first peak in the DTG.
step 3	wt %	Expressed as a percentage of the total sample weight. Loss of weight in the area of the second peak in the DTG.

4.3 Results

4.3.1 Thermo-gravimetric analysis

4.3.1.1 Pyrolytic behaviour of the biomass

Thermogravimetric analysis and a differential scanning calorimetry (TG-DSC) in an inert atmosphere (Ar) were carried out simultaneously on the samples to evaluate the pyrolytic behaviour of the biomass. The information returned by the analysis was the amount of weight loss and the heat exchanged by each sample during the run.

The DTG curve describes the rate at which the mass loss happens as the temperature increase and provides information about the mass loss. TGA allowed the identification of three main regions of weight loss (Figure 4-1). The first, between 40 and 105 ℃, corresponding to the loss of moisture from the biomass, the second between 250 and 350 ℃ and the third between 350 and 450 ℃. The second and the third region where the phases where the maximum loss of mass happened and they are due to loss of

volatile substances, which for lignocellulosic material is usually a biphasic process correlating to the volatilisation of cellulose and hemicellulose. Considering the three main areas identified, some variates, able to quantify the pyrolytic behaviour in each region were obtained. A list of the variables and their description can be found in Table 4-1. For all the three areas, the amount of mass lost was calculated. Also, for the region of maximum weight loss, the onset and end set temperature were identified. The experiment was performed in two technical replicates on the three biological replicates of each genotype. Values for each of these variates were averaged between the three biological replicates of each genotype and genetic group and expressed as mean±sd (Table E. 1). The statistical significance of the differences was investigated using an analysis of variance (ANOVA) (Fisher, 1919) and further localised using a Tukey HSD test (Tukey, 1949). Results are presented in Table E. 2. The onset temperature varied between 284.21 and 323.29 ℃, and while the endset temperatures varied between 356.18 and 386. Although significant differences were detected between genotypes by ANOVA, Tukey HSD test did not detect statistically discrete groups between the genotypes analysed (Figure 4-3). This is usually the case when there are many pairwise comparisons (which is likely the case when investigating an interaction term). The Tukey procedure controls the Type I error rate and requires a larger difference to declare significance compared to if no adjustment was used. The ANOVA F-test uses mean standard error (MSE) in the denominator which borrows information from all the data and is not affected by this adjustment.

Interestingly the two variates have the same trend across the group of genotypes, indicating that lower onset temperature corresponded to lower endset and vice versa. The relation between the two variables is investigated in more detail below (subsection 4.3.1.2). As much as the mass loss in the different steps is regarded, there was a

significant difference between genotypes only in the mass loss in step 3. However, no genotypes were utterly distinct from the others.

Conversely, the differences between genetic groups were statistically significant (Figure 4-4). In terms of onset and endset temperatures and mass loss in step 3, *M. floridulus* was not significantly different from the artificial hybrid. At the same time, it was significantly different from the other two hybrid groups. For all these three variates, *M. lutarioriparius* was significantly different from *M. sacchariflorus*. The four groups of *M. sinensis* were not differing from each other and the rest of the group for any of the variates.

Overall, these results show a limited but significant difference between genotypes and genetic groups in terms of onset and endset temperatures and the amount of mass loss between 350 and 450 °C.

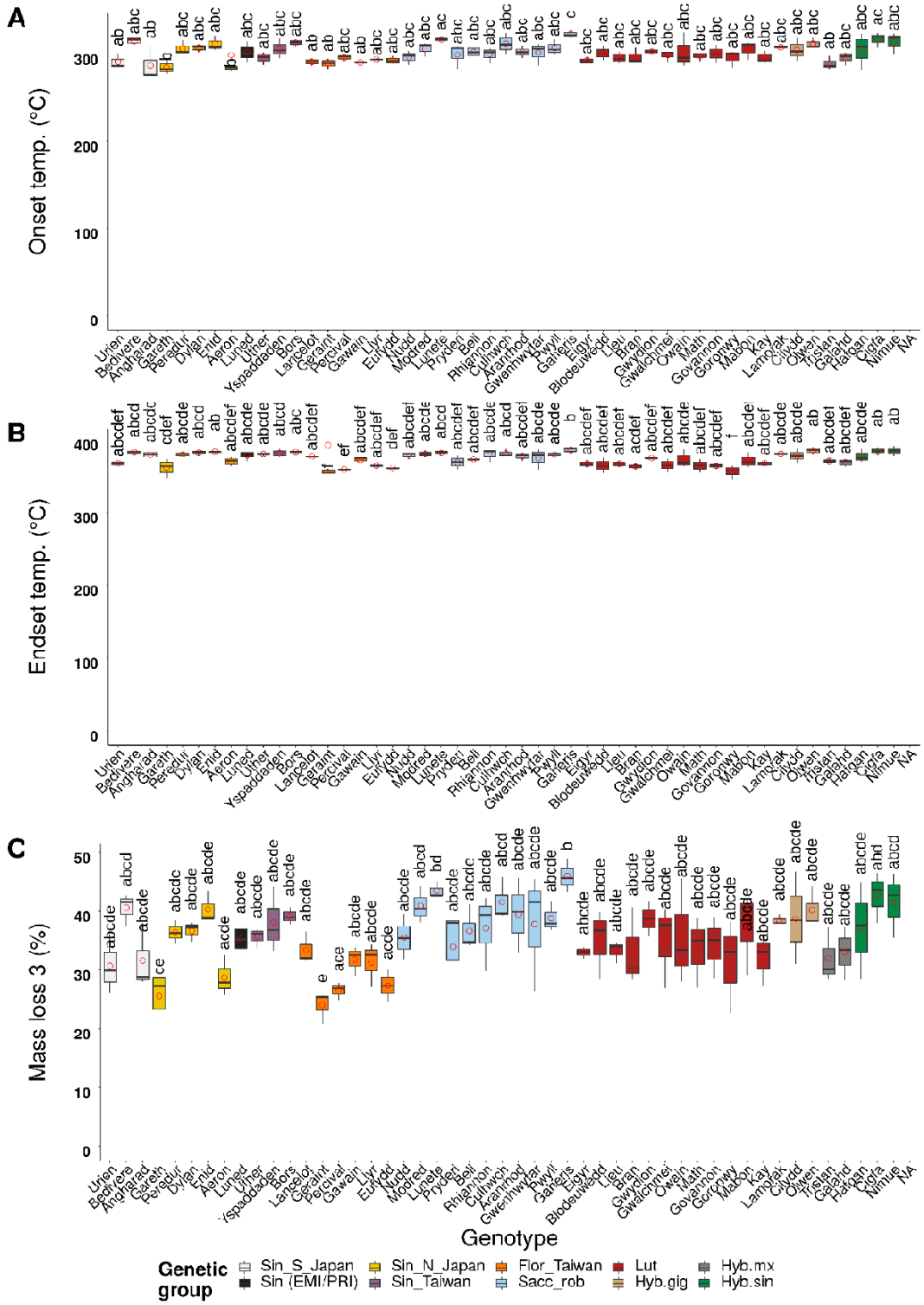


Figure 4-3 Effect of genotype on three parameters of the thermogravimetric curve

Box and whiskers plot showing the significant effect of the genotype from which the biomass was collected on the onset temperature (A), endset temperature (B) and mass loss between 350 and 450°C (C) The significance was assessed using a Type II ANOVA ($p<0.01$). Pairwise comparisons were performed using a Tukey test for significance with $p=0.05$ and represented by the letter labels on the top of each box.

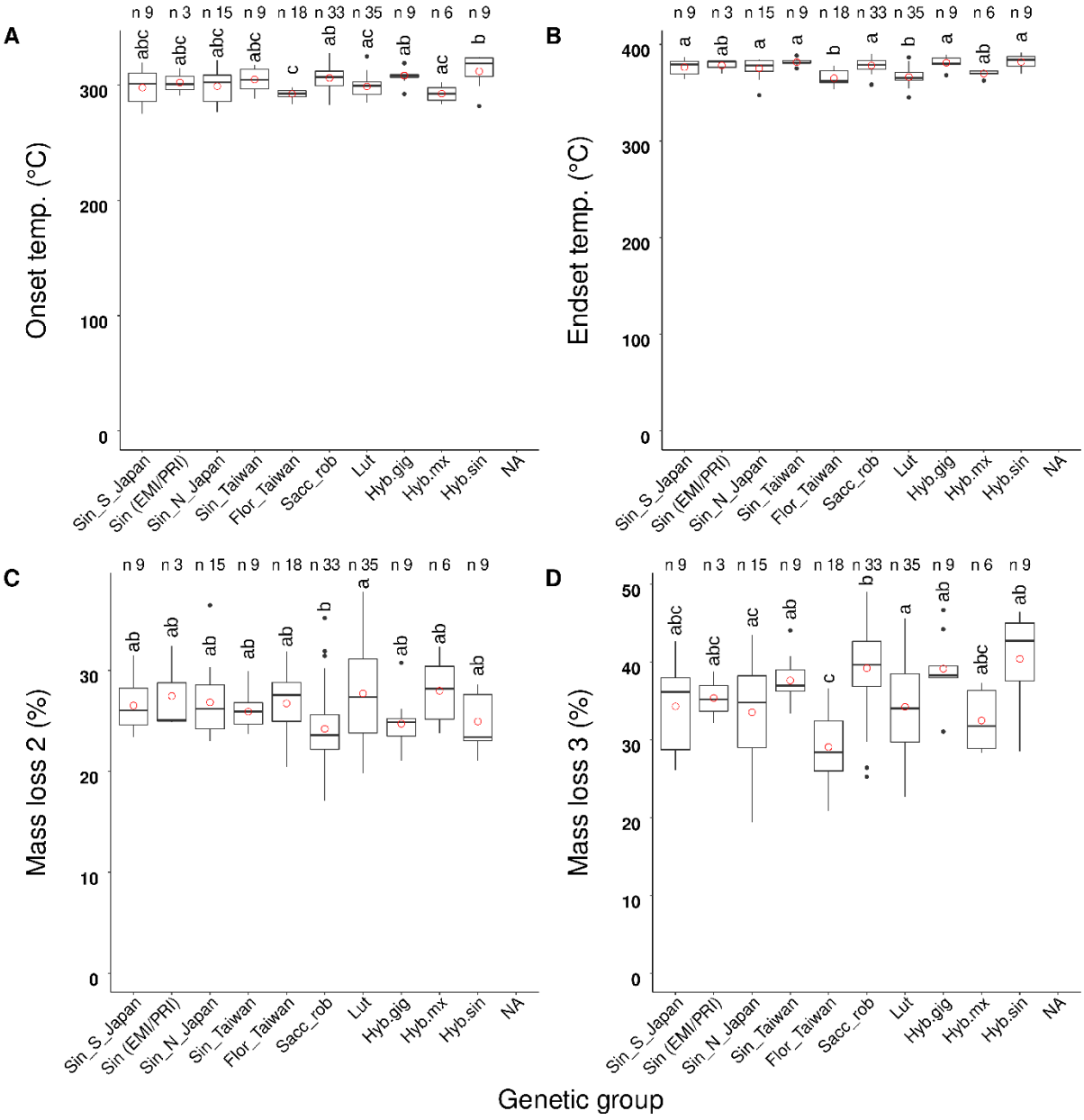


Figure 4-4 Effect of genetic group on four of the parameters of the thrmogravimetric curve

Box and whiskers plot showing the significant effect of the genetic group ($p<0.001$) on four parameters of the mass loss curve obtained from the thermogravimetric analysis (TGA) of the biomass from the 49 genotypes. The significance was assessed using a Type II ANOVA test and the pairwise comparison between the genetic groups was performed using a Tukey High Significance test ($p=0.05$). Significant pairwise comparisons are indicated by different letters on top of the boxes. Labels on top of the boxes indicate the number of observations per genotype. Each genotype was measured in one technical replicate and three biological replicates. A. Effect of genetic group on onset temperature for the maximum weight loss region during the TG. B. Effect of genetic group on the endset temperature for the maximum weight loss region in the TGA curve. C. Mass loss, expressed as %w/w, between 250 and 350 ℃ in the TGA curve corresponding to the region of the first peak in the differential thermogravimetric (DTG) curve. D. Mass loss expressed as %w/between 350 and 450 ℃ in the TGA curve corresponding to the second peak in the DTG curve.

4.3.1.2 Correlation with the biomass characteristics

The relationships between the thermogravimetric variates and the cell wall composition and saccharification efficiency presented in chapter 3 were tested for correlation using the Pearson method between onset temperature, endset temperature, mass loss in the three steps and all the variates of the biomass was tested. Only correlation with an r-square value higher than 0.45 was considered relevant and will be presented and discussed. For each relevant association, the significance of the relationship was tested and given in a scatter plot, including the main statistics (Appendix 4.3). As expected from the observation of the summary statistics, there was a strong positive correlation between onset and endset temperatures ($R = 0.7$, $p < 0.0001$).

Endset and onset temperatures both correlated negatively with mass loss during the first step of the maximum loss of weight and positively with the second step. In other words, the higher temperature at which the maximum mass loss started was the one that corresponded to the higher amounts of biomass loss between 250 and 350 ℃ and a more elevated amount of mass loss between 350 and 450 ℃. There was a strong positive correlation between onset temperature and endset temperature with cellulose content, but not with hemicellulose. Moreover, cellulose content correlates well only with the mass loss in the second step of maximum mass loss. These results suggest that onset temperature could predict the amount of volatiles produced from the biomass. At the same time, the mass loss in the second step of pyrolytic degradation could estimate the amount of cellulose initially present in the biomass.

Some other parameters of biomass composition were well correlated to thermogravimetric measurements. The onset temperature was relating positively to NDF, ADF, ADL and Klason lignin. Among them, the strongest correlation was with

NDF (R = 0.68). The endset temperature related positively to NDF and ash. Onset temperature also correlated negatively with the amount of glucose enzymatically released as a percentage of CWM (Glc_Enz_p). Endset temperature related negatively to the amount of xylose enzymatically released by a unit of weight of CWM (Xyl_Enz_p). Overall, these results show that the onset and endset temperatures can provide useful information about the biomass composition and quality for other conversion processes. The increase of NDF, decrease of ash, and increased amount of xylose released was correlated to lower endset temperatures.

The amount of mass loss in step 3 correlated positively with NDF, ADF, ADL and Klason lignin contents and negatively to ash content. The minerals in the ash are not volatilised, and a higher amount of ash corresponds to a lower peak.

Finally, correlations were detected between thermogravimetric parameters and structural characteristics of the cell wall. The ratio between holocellulose and lignin (H:L) negatively correlated with the onset temperatures and with the mass loss in step 3. This result indicates that H:L can determine the temperature at which volatilisation of material begins and that the lignin has a peak of degradation around 450 ℃ and is, therefore, contributing mostly to the second peak in the area of maximum mass loss.

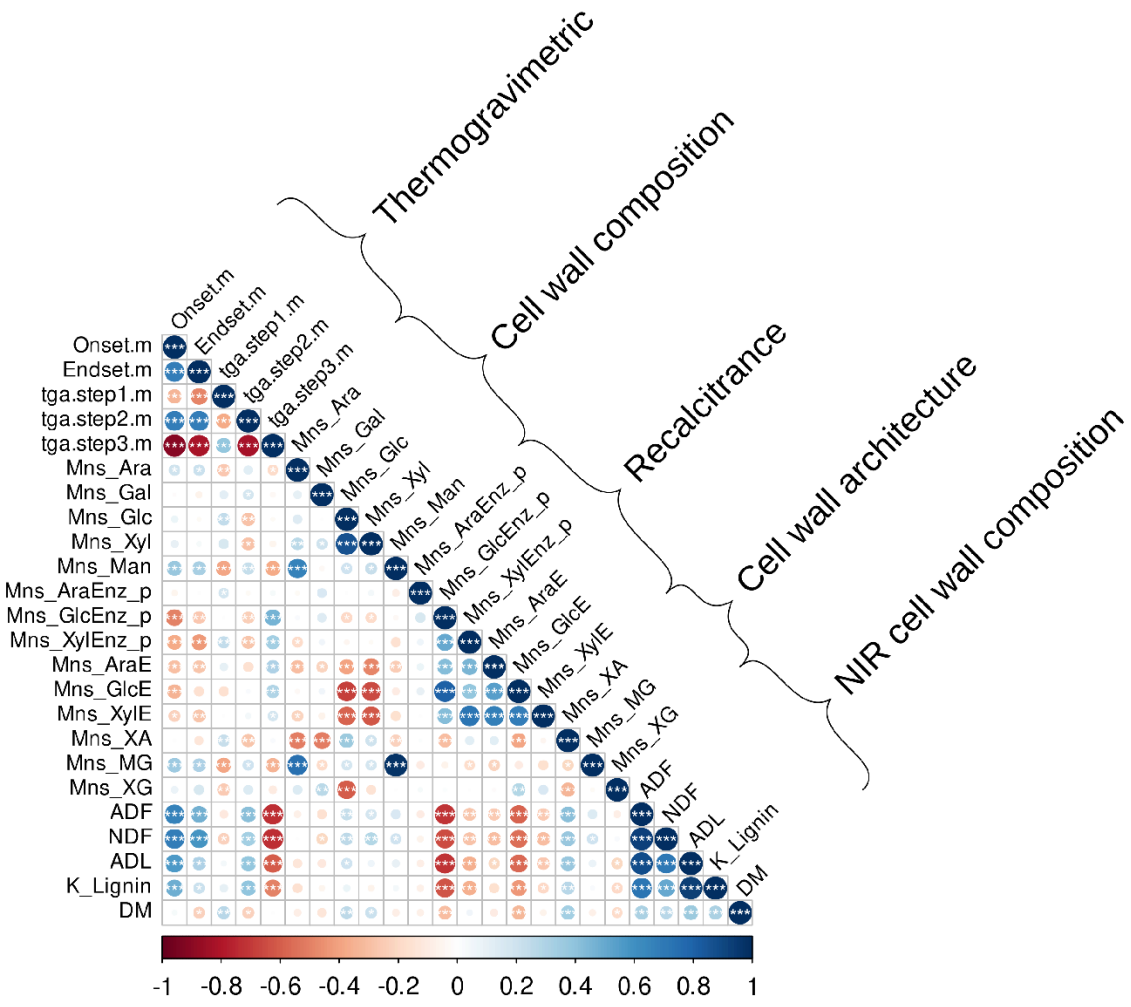


Figure 4-5 Correlation between parameters of the thermogravimetric curve and cell wall traits.

Correlation plot representing the coefficients of regression calculated using a Pearson method between the parameters obtained in the thermogravimetric analysis (TGA) (Thermogravimetric), the variables measuring the cell wall composition, cell wall recalcitrance, cell wall architecture and cell wall composition measured by nuclear infrared (NIR) spectroscopy. The colour scale indicates the R value for the correlation and varies from 1 (blue) to -1 (red). The correlations were tested for significance and p values were corrected for multiple comparisons using the Bonferroni method. Stars indicate the threshold of the adjusted p value. Adj. p: “***” .001, “**” .01, “*” .05. Non-significant correlations are blank.

4.3.2 Principal least square regression of TGA/DSC data

Next, the ability to predict biomass traits from TGA/DSC data was tested. A principal least square regression (PLSR) approach was used to find a regression between the TGA/DSC curve and the variates determined in chapter 3. The multivariate approach was chosen because of its ability to rapidly predict biomass characteristics has been reported (Acquah et al., 2017).

The model was created to predict the amount of glucose (GlcE), and xylose (XylE) enzymatically released, the total content of glucose (Glc) and xylose (Xyl) in the cell wall from the TGA and DSC spectra. The prediction plots for each model are presented inFigure 4-6 A-H. For the GlcE a PLSR model was created with 11 components both for TGA and DSC. The cross-validation resulted in an R2 of 0.67515 for the TG data and 0.5207 for the DSC data. For the xylose enzymatically released (XylE), the prediction was possible using 19 components for the TGA and 8 components for the DSC dataset. The model had an R2 of 0.62713 for the TG data and 0.53019 for the DSC data. The amount of glucose in the cell wall was predicted by a PLS with 19 components on the TG data and 17 components on the DSC data. Finally, the PSLR model for the amount of xylose in the cell wall material (Xyl) with 15 components of the TG and 19 components of the DSC had an R square of 0.54388 and 0.636003 respectively. All the models were able to explain more than 95% of the total variance of the dataset.

The plot of the loadings (Figure 4-7) of the first 4 components of the PLSR model of TG data for the four traits studied showed characteristic peaks in the region between 960-1560 seconds. The time corresponds to the region of maximum weight loss. The shape of the loading curve was similar between Glc and Xyl and between XylE and GlcE. In contrast, it was different between the two pairs of variables. From this result,

it could be hypothesised that the pattern of the loading curve could inform on the nature of the variable.

These results show that it would be possible to predict biomass quality and composition traits from TG and DSC spectra.

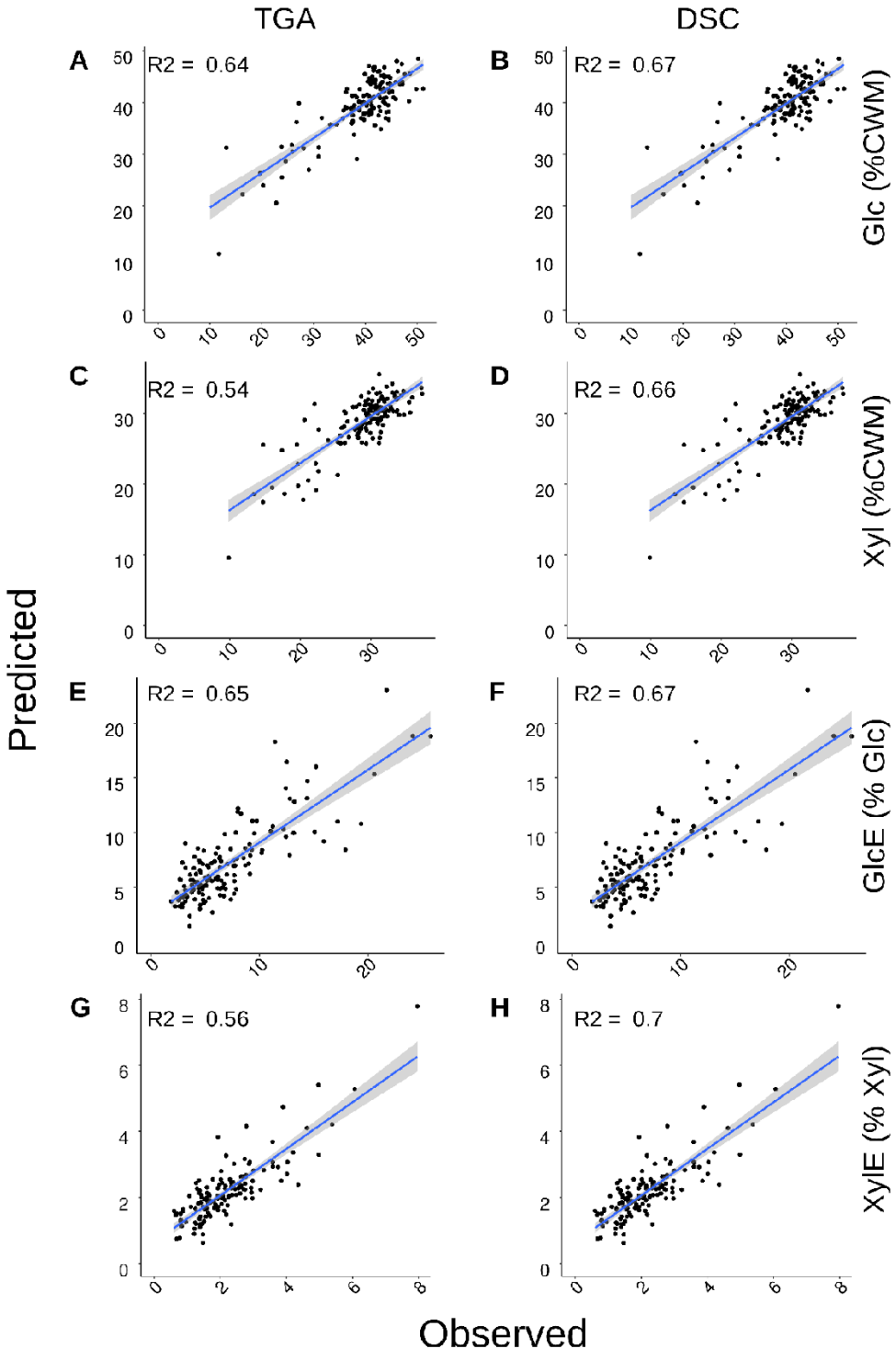


Figure 4-6 Regression plots between observed and predicted values of four cell wall variables.

Fit between values for cell wall characteristics observed and predicted by the principal least square regression (PLSR) of the thermogravimetric analysis dataset (TGA) and of the differential scanning calorimetry (DSC) dataset. The cell wall characteristics considered were the glucose content (Glc) (A, B), the xylose content (Xyl) (C, D), the amount of glucose enzymatically released (GlcE) (E, F) and the amount of xylose enzymatically released (XylE) (G, H).

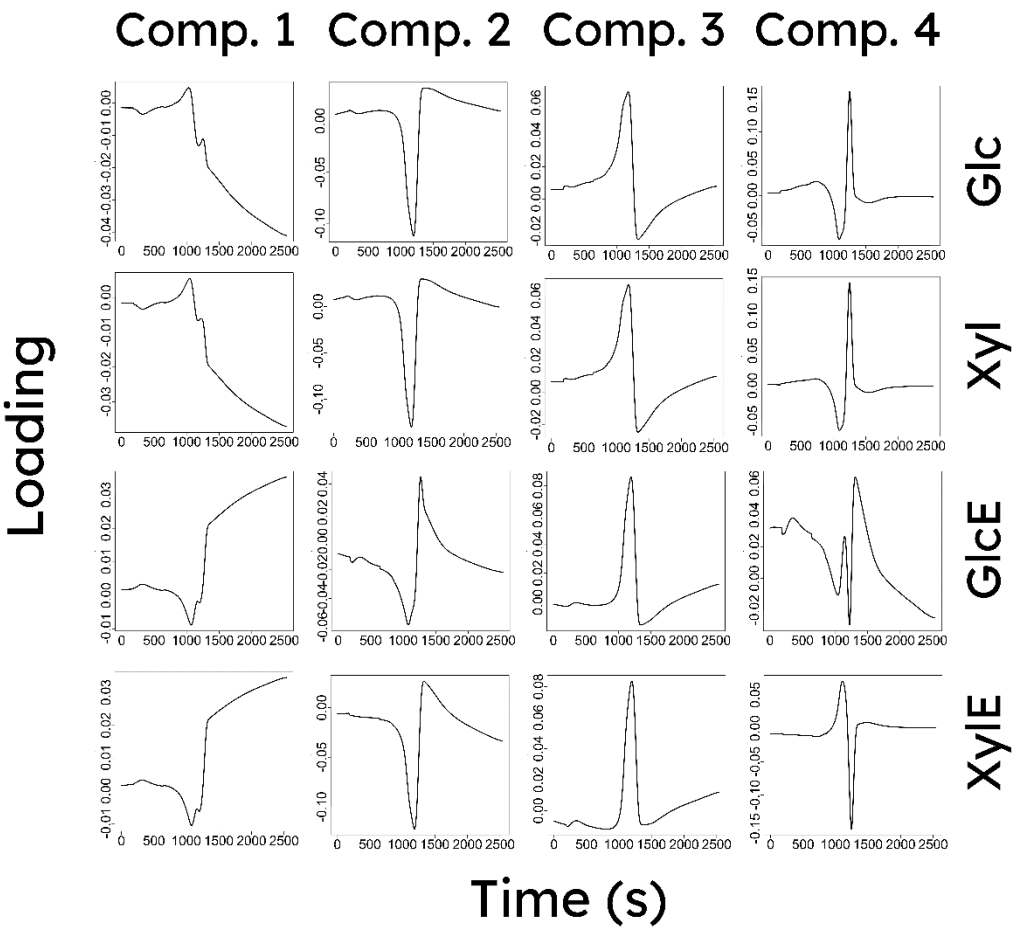


Figure 4-7 Loadings of the first 4 components of the principal least square regression (PLSR) models

The thermogravimetric profile could be used to predict the recalcitrance of the biomass to enzymatic deconstruction. Here a plot of the loadings of the PLSR models using thermogravimetric dataset to predict the amount of glucose (Glc) and xylose (Xyl) in the cell wall and glucose (GlcE) and xylose (XylE) enzymatically released from the cell wall. It shows which timeframes (x axis) affect the first 4 components used for the model.

4.3.3 Evolved Gases Fourier Transformed Infrared Spectroscopy (EG-FTIR)

The complete set of data obtained from the EG-FTIR of the 119 samples can be imagined as a 3 – way matrix created by three bidimensional objects:

$$x(i,j,k) = a(i)b(j)c(k)$$

Where the first dimension, *i,* is represented by the wavenumber (from 2000 to 5000), the second dimension, *j* , is the elution time (12 to 22) and the third dimension, *k,* is the samples (1 to 119). Each of these objects had a fourth dimension ® represented by the absorbance at each wavenumber.

The dataset was analysed using a PARAFAC method. PARAFAC is a decomposition method, is *one* generalization of bilinear PCA, and allowed to find another array, that best represents the data.

Two different types of PARAFAC have been developed PARAFAC and PARAFAC 2. PARAFAC 2 was developed to solve the low level of fit obtained from time-shifted spectral data. In the case of FTIR of evolved gases from TGA, since there was no monitoring of the actual flow of gas at the end of the system, it was considered convenient to check the spectra alignment for the occurrence of peak shifts. The shifting was tested visually by plotting the spectra of all the genotypes at each time point together. From Figure E. 1, it is possible to see that there was no visible shift in the peaks at any time point. From the observation of the 3D image of the absorbances for each sample, it was possible to observe a negative peak in the area before 2000 nm cm^{-1}. After comparison with the spectrum of a blank sample, it was deduced that the negative peak was not backgrounded noise. CO2 usually causes peaks in that region. Although the reason for the negative peaks was not found, the data used for the PARAFAC model were limited to the area between 2000 and 5000 cm^{-1}.

The optimal number of components to retain in the CP model was calculated using Core Consistency (CORCOND) method suggested by Bro and Kiers (2003) implemented in the R function `multiway::corcondia()`. The result of the test is shown in Table 4-2. The value of the quality of the fit of the model decreased in the first three components, and this showed that three components were appropriate for the analysis.

Table 4-2 Core consistencies according to CORCONDIA

Core consistency test versus the number of components in an unconstrained PARAFAC model of the FTIR data.

Component	CORCON
1	100
2	89.13527
3	64.75288
4	-182.272

Accordingly, a CP (PARAFAC/CODECOMP) with three components was fit to the three-dimension dataset. For each component in each modality of the model, the loadings were plotted. The model was able to detect an underlying structure with three modalities in the dataset that is evident in its graphical representation. The first modality represents the absorbances at the different wavenumbers. From its loading plots, it possible to see that four main peaks can describe the structure of the data. The first between 2000 nm cm^{-1} and 2500 nm cm^{-1}, the second between 2740 and 2984 nm cm^{-1} peaking at 2919 nm cm^{-1}, the third between 3200 and 3167 nm cm^{-1} peaking at 3069 nm cm^{-1} and a fourth, higher one between 3190 and 3522 nm cm^{-1} peaking at 3366 and 3440 nm cm^{-1}. The last peak could also be interpreted as a double peak. All

the four peaks predicted by the model can be observed in different extent in the set of samples.

In the first modality, the first factor shows only the fourth peak with a negative effect. The second and third factor shows all the four peaks. Still, they differ in the direction of the loading in the sense that the second factor is negative as the first, while the third factor is positive.

The second modality is describing the structure of the dataset concerning the time of collection of the spectrum. In Figure 4-9, the x-axis shows the time from 1 to 34. Each point corresponds to data acquisition from the FTIR along the 40 minutes run of the TGA. Therefore, each factor on this modality is describing when the spectrum observed in the first modality was emitted. It is possible to see three peaks in the time. Aligning the time modality loading plots with the TGA program, the three peaks correspond to the three main areas observed in the TGA. The first loss of weight due to moisture loss, the main degradation area where most of the mass is loss by volatilization and the final decomposition area. The time loadings were negative for all the three factors. In the first factor, the central peak is the most prominent, in the second factor, the first peak is the most prominent, and finally, in the third factor, the three peaks have comparable sizes. This result indicates the emission of different classes of compounds volatilised during the run. The higher number of peaks in the second factor indicates a higher number of complex molecules volatilised during the maximum mass loss.

In the sample modality, the three factors are grouping the samples differently. Three groups can be seen along the first factor, five along with the second and six along with the third factor. When coloured according to the saccharification efficiency as determined in chapter 4, one of the groups in all the three factors is represented by the

genotypes with the highest enzymatic sugar release. After testing for the other dependent variables measured previously, it was not possible to identify the grouping factor able to explain all the groups in the three factors. Taken as a whole, these can suggest that FTIR of emitted gases could be developed as a technique for the assessment of recalcitrance traits in biomass.

Background noise at the beginning and the end of each set of measurements for each sample could affect the model. Consequently, a new model was created, considering only the spectra acquired during the maximum loss of weight (time points from 11 to 22). Since the time selection was arbitrary and to avoid that time could affect the model in this second case, a PARAFAC2 was used to fit the data.

The three-dimensional representation of the sample modality of the dataset (Mode C of the CP) (Figure 3.17) shows that the group of genotypes with the highest saccharification efficiency is separating slightly from the rest of the samples.

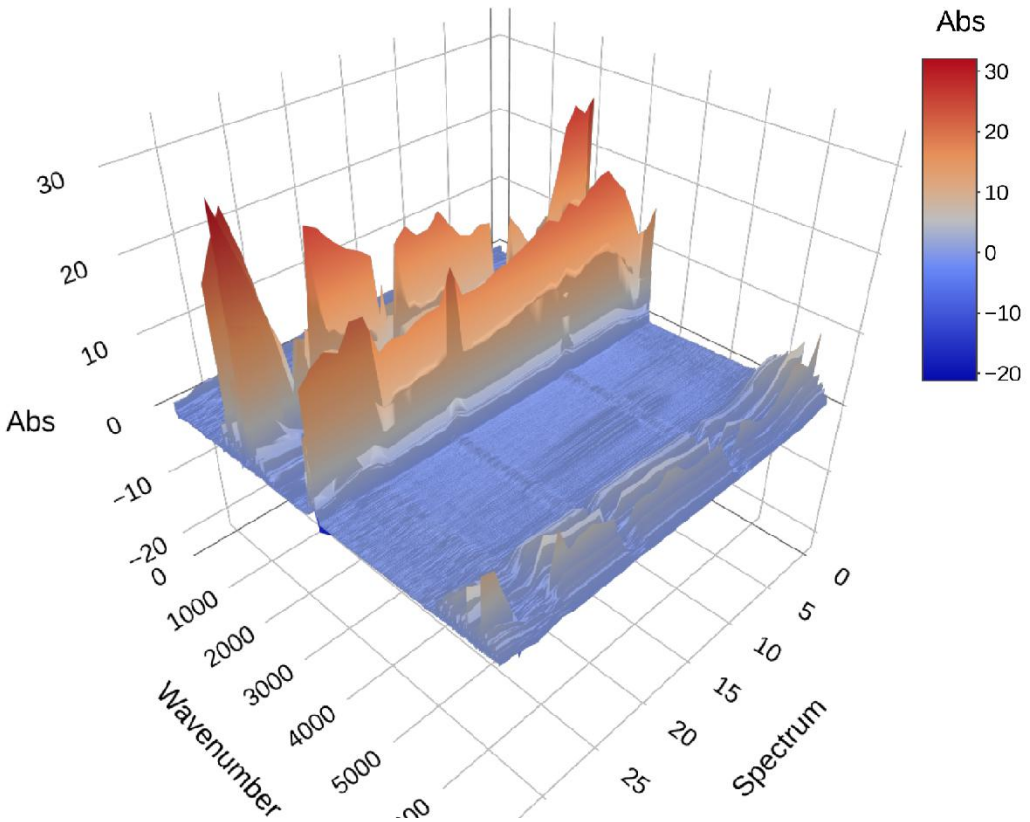


Figure 4-8 Example of Fourier transformed infrared spectroscopy (FTIR) spectrum of evolved gas

Graphical representation of the dataset resulting from FTIR analysis of the evolved gases by the thermogravimetric analysis (TGA) on the biomass from genotype 19, replicate 1. The dataset is tridimensional and records the value of absorbance (Abs on y-axis) on 34 spectra collected during the TGA running time (Spectrum, z-axis) at different wavenumbers (x-axis).

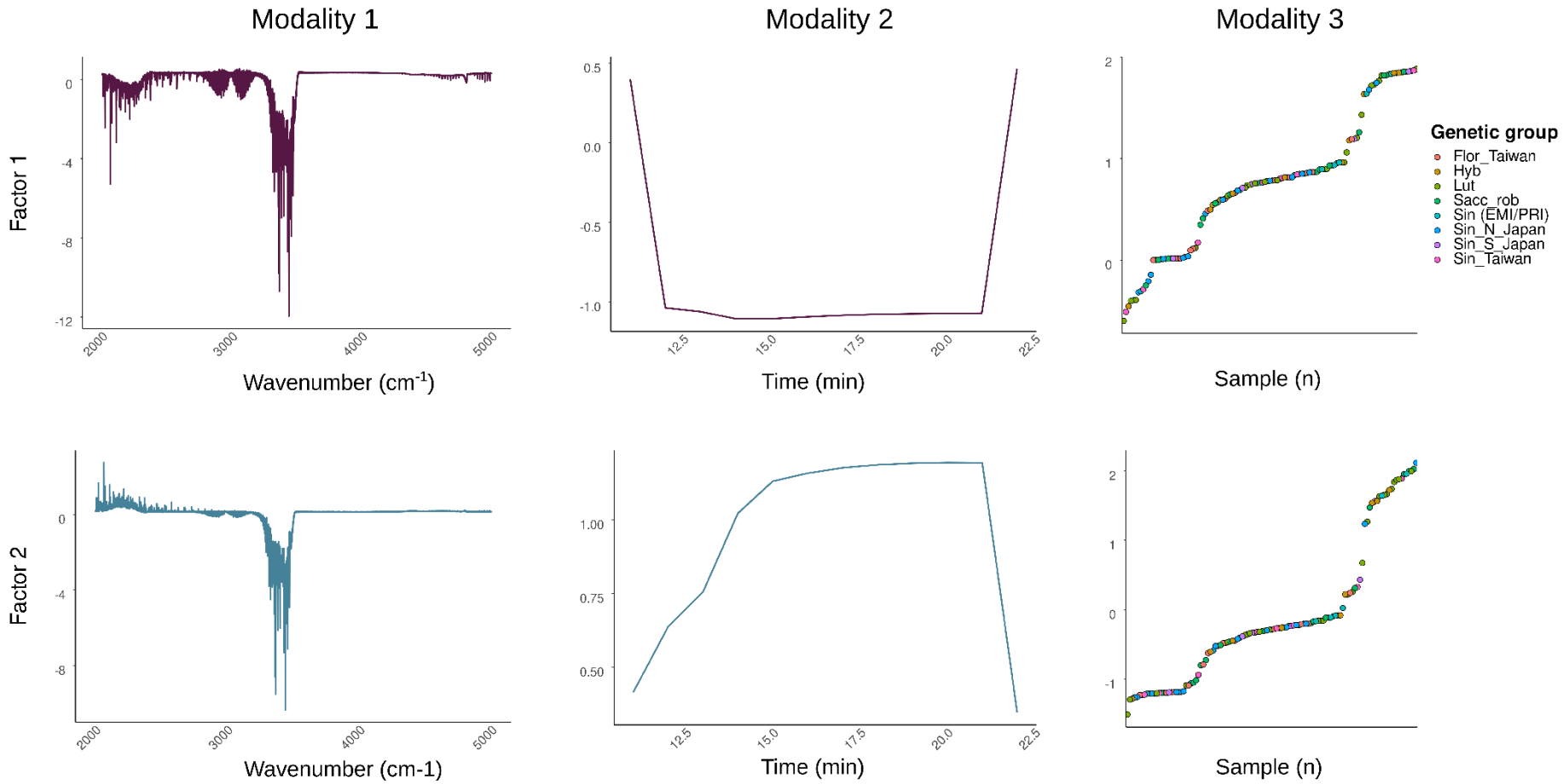


Figure 4-9 PARAFAC. Modalities and factors
The output of the unconstrained PARAFAC test. The loadings of each factor (x-axis) are plotted in the three modalities. The first modality (wavenumber) identified the peaks that are mostly contributing to the variance. The second modality (time) describes the variation of the peak abundance in time, the third modality (sample) describes the loading for each sample. The third modality created a separation between samples, but this was not associated with any of the traits studied. Colour of the point in the third modality represents the genetic group.

4.4 Discussion

4.4.1 Pyrolytic behaviour of *Miscanthus* biomass

The pyrolysis of the biomass was investigated using a TG-FTIR system. The use of TGA as a fast technique for proximate analysis of plant biomass is well established (Saldarriaga et al., 2015). The thermogravimetric study provided information about the pyrolytic degradation of the biomass. Three steps were identified, an initial loss of weight-related to the loss of moisture and two subsequent steps determining the maximum weight loss. Similar behaviour was reported for other types of biomass. Wang et al. (2013) identified three stages in the degradation of poplar biomass. The first stage, below 180 °C, characterized by a mass loss for drying, the second stage, between 180–260 °C where mass loss was due to a small amount of depolymerization and recombination of internal material and a third step, between 260–380 °C, where the maximum and rapid mass loss happened and CO_2 and some volatile organic compounds, such as carboxylic acid and aldehyde, were released along with small amounts of NH_3 and CO. Lu et al. (2009), reported a similar behaviour in rice straw. Although they indicated the process in four stages, the first and second can be assimilated to the one stated by Wang et al. (2013) as first.

There was a limited but significant difference between genotypes and genetic groups in terms of onset and endset temperatures and in the amount of mass loss between 350 and 450 ℃. This could depend on the rearrangement of the bonds during the previous step making the compositional differences not have an effect in this step. Furthermore, the onset temperature was relating to the amount of volatiles produced that during the second step and with the amount of cellulose initially present in the biomass. This is in agreement with the fact that at the heating rate used in this experiment (25℃/min) the volatilisation of heavy volatiles has been reported for

Miscanthus (Kok and Özgür, 2013). The use of integration of DTG peaks has been proposed as an effective and economically viable method for the assessment of biomass quality parameters (Sher et al., 2019).

Moreover, results show that the onset and endset temperatures provided information about the biomass composition and recalcitrance to biochemical conversion. The higher NDF, lower ash, and higher xylose released were all correlated to lower endset temperatures. This is the first time that thermogravimetric parameters have been associated with complex quality traits such as the amount of glucose or xylose enzymatically released.

4.4.2 Prediction of biomass traits from TGA and DSC data

Fast and high throughput methods are necessary to assess the quality and composition of the biomass. Here we showed that a TG analysis coupled with a DSC can provide a fast and reliable tool for predicting the recalcitrance of the cell wall to deconstruction and the cell wall composition. The number of components used to build the models was able to explain more than 95% of the total variance in the dataset.

Acquah et al. (2017) found that a statistical approach using PLSR can be used to determine the chemical composition of the biomass starting from TG data. Also, they showed that PLSR models had a lower cross-validation error than traditional kinetic models. Traditional kinetic models based on the deconvolution of the Differential Thermogravimetric Curve (DTG) (Marini et al., 1979; Zhang et al., 2014b).

4.4.3 PARAFAC of EG-FTIR

The PARAFAC approach allowed to build a unique trilinear model of the spectral data collected. This approach allowed to interpret a large number of spectra of complex samples and find a lower-dimensional model able to explain the variability in the spectra. It has been rarely used before.

The PARAFAC analysis of the EG-FTIR revealed three main peaks in the three factors of the first modality (Figure 4-9). This was probably a consequence of the emission of different classes of compounds that volatilised at different temperatures. According to the higher number of peaks in the second factor of the first modality, the compounds emitted in the area of maximum weight loss were more complicated and consequently with a higher molecular weight. Pyrolytic products of lower molecular weight feedstocks have an increased C/O ratio compared with cellulose (Mettler et al., 2012). They showed that low molecular weight feedstocks generate 2–3 times as many furans (such as hydroxymethylfurfural, HMF) as cellulose (Mettler et al., 2012; Wang et al., 2013b). The production of lower molecular weight feedstocks has been associated with the level of crystallinity of cellulose and to the milling process (Wang et al., 2013b). Considering that HMF is a vital platform chemical for the bioeconomy, this observation could be practically applied to the characterisation of biomasses for the production of chemicals.

Moreover, the separation of genotypes with different levels of recalcitrance to saccharification along the third modality of the PARAFAC suggested that EG-FTIR could be developed as a technique for the assessment of biomass quality for biochemical conversion. This would accelerate the process of selection of germplasm with superior characteristics in terms of biomass recalcitrance to conversion and the production of plant-based chemicals.

5 Effect of organ on the composition of the cell wall of *Miscanthus* genotypes

5.1 Introduction

Cell wall structure and composition are related to the age of the cell and, in turn, with the developmental stage of the plant. In the apical meristem, just after cell division, plant cells are surrounded by a thin layer of pectin rich cell wall, the middle lamella. As the cell matures and extends new layers of material are juxtaposed on the inner face of the middle lamella creating a cellulose and hemicellulose rich layer, partly extensible, the primary cell wall. The primary cell wall is rigid enough to control the direction of cell extension, though flexible enough to adapt to the increase in the size of the cell. Compositionally, the primary cell wall is composed of cellulose embedded in a matrix of cross-linking glycans and proteins. Once the cell elongation has ceased, new material is added to the cell wall to create a secondary cell wall, resulting from a complex of genetic and environmental stimuli. In many types of cells, it is the secondary cell wall that gives the cell the ability to specialise. The main difference in secondary walls is the addition of lignin (Albersheim et al., 2011; Alberts et al., 2002; Fangel et al., 2012) .The same can be observed on a larger scale for tissues and organs. It is therefore not surprising that the composition of the biomass, mostly made of cell walls, is a dynamic trait, changing with the developmental stage of the plant .

Therefore, in investigating the quality of biomass for conversion, the influence of developmental stage and organ must not be forgotten. Developmental related differences in cell wall have been investigated both in dicotyledons and monocotyledons (Cosgrove et al., 2012; Hatfield et al., 2017). Differences related to

organ and developmental stage in cell wall composition and cell wall component interlinkage were assessed on the model *Brachypodium dystachion* and in the main annual monocotyledons (Christensen et al., 2010; Rancour et al., 2012). More recently, non-model monocotyledons have been considered. Perrier et al. (2017) reported differences in cell wall composition in sorghum between different plant portions (node vs internode) and between tissues in the same part (inner internode vs outer internodes). Also, in *Miscanthus,* differences in the cell wall composition and structure between plant portions and organs have been reported (Hodgson et al., 2010; da Costa et al., 2014). In *Miscanthus*, Le Ngoc Huyen et al. (2010) reported a lower arabinose ramification in the internodes with a value of the xylose-to-arabinose ratio three times higher in that part of the plant, as compared to leaves and sheaths, indicating a higher proportion of primary cell walls in the latter organs. Also, stem material of *Miscanthus* has a higher lignin content when compared to leaf material (da Costa et al., 2014). The amount was higher in the basal internodes compared to the more apical ones (Le Ngoc Huyen et al., 2010). Moreover, the linkages in lignin were differing between leaf and stem. The lignin in the stem material contained a higher proportion of monolignols and had a higher ratio between syryngyl and guaiacyl units (Le Ngoc Huyen et al., 2010).

Differences in biomass composition between leaf and stem material of *Miscanthus* result in differences in the conversion quality between the two organs. In terms of quality for biochemical conversion, leaf biomass is usually less recalcitrant than stem biomass (Le Ngoc Huyen et al., 2010).

The selection of the 12 genotypes to be used in for the nutrient stress experiments, described in Chapter 6, was based on the assessment of their biomass quality from pooled aboveground biomass samples. Considering that the material was harvested

from plants of the same age, it was hypothesised that some of the differences observed in cell wall composition and saccharification efficiency between genotypes and genetic groups, found in Chapter 2, could have been a consequence of differences in the ratio between organs whose biomass had a different composition. Here the differences in the cell wall structure and composition in the 12 *Miscanthus* genotypes selected in chapter 3 are investigated for leaf and stem material.

5.2 Material and methods

5.2.1 Plant material

In January 2017 biomass material from the senesced plant of the 12 genotypes selected in the previous chapter was collected from ABR33. From each plant, three full tillers were collected. The material was separated into leaf (leaf blade and leaf sheet) and stem. The material was chopped in smaller pieces and subsamples were collected in plastic bags before being freeze-dried. Weight of the whole fresh material and subsamples before and after freeze-drying was recorded for both portions. The difference in weight between the fresh and dried weight of the subsamples was used to determine the moisture of the whole plant, similar to what was done by Clifton-Brown et al. (2019).

5.2.2 Plant pictures

In January 2018 pictures of the whole plant grown in ABR33 for each of the 12 selected genotypes were taken using a DSLR Canon camera. Variable weather conditions and time limiting situation did not allow to have standardised pictures that could be further analysed. However, they have been used for informative purposes.

5.2.3 Leaf stem ration

The calculated dry weight of leaf and stem biomass was used for the calculation of the leaf stem ratio for each genotype according to methods used in literature (Smart et al., 1998; Zhang et al., 2014a).

5.2.4 Biomass grinding

Biomass material was ground using a hammer miller with a 1mm mesh. Further milling was performed according to that described in 3.2.1.2.

5.2.5 Cell wall preparation

5.2.5.1 Total solid determination

Total reliable determination of solids in the biomass was carried out according to (Sluiter et al., 2008). Eppendorf microcentrifuge tubes were dried by placing them in a 105 ± 3°C drying oven for four hours and weighted to the nearest 0.1 mg. Weight was recorded. The milled sample was thoroughly mixed, and approximately 2 mg were weighed out to the nearest 0.1 mg, into the tubes. Weight of the sample plus the tube was recorded. The experiment was performed in duplicate. The sample was placed into a convection oven at 105 ± 3°C for four hours, after that the weight of the sample plus tube was re-recorded. Then samples were put into a convection oven at 105 ± 3°C until constant weight. Weight was recorded 24 and 48 hours after the transfer.

5.2.5.2 Cell wall purification

Cell wall material was prepared as described in subsection 2.2.1 and 3.2.1.3.

5.2.5.3 Determination of the total cell wall monosaccharide content

The total amount of monosaccharides contained in the cell wall was quantified as explained in subsection 3.2.1.4.

5.2.5.4 Hemicellulose composition

Hemicellulose was isolated according to the protocol published initially by Albersheim et al. (1967) as reported by Foster et al. (2010b) and its composition in terms of monosaccharides assessed as described by (da Costa et al., 2017; Petit et al., 2019).

To perform the method, 2 mg of cell wall material was weighed into 2ml Sarstedt tubes by hand. Weight was recorded to the nearest 0.1mg. The walls of each tube were rinsed with 250 µl of acetone to collect the cell wall material on the bottom of the tube, and acetone was evaporated under very gentle airflow. Then 250 µl of 2M trifluoroacetic acid (TFA) was added to each sample. TFA was added carefully to ensure no material was splashed up onto the tube walls. Tubes were capped tightly and incubated for 90 min at 121°C in a heating block. After this time, the heating blocks and samples were cooled on ice and centrifuged at 10,000 rpm for 10 min. 100 µl of the acidic supernatant containing the matrix polysaccharide derived monosaccharides were transferred to a clean Eppendorf tube, making sure not to disturb the pellet material. Monosaccharides in the solution were separated and quantified by High-Performance Anion Exchange Chromatography and Pulsed Amperometry Detection (HPAEC-PAD) respectively using a Dionex ICS5000® from Fisher Scientific operated at 45°C using a CarboPac SA10 (4×250mm) column with a CarboPac SA10G (4×50mm) guard column. An eluent generator coupled to the system continuously prepared a KOH solution at 0.001M for isocratic elution at a flow rate of 1.5cm^3/min for 14min. In all cases, a volume of 0.025cm^3 of the sample was injected into the column and detected by PAD using a gold working electrode and an Ag/AgCl reference electrode. Samples were diluted accordingly to fit the calibration standard curve.

5.2.6 Lignin content

The Acetyl bromide soluble lignin method is a rapid and reliable procedure for lignin quantification in small biomass samples. The method for determination of lignin with acetyl bromide described below was initially published by (Johnson, 1961) and modified by (Fukushima and Hatfield, 2001; Foster et al., 2010). Recently, the protocol was modified (upscaled) by (Costa, 2015) to reduce the weighing error due to electrostatic repulsion.

The method is reliable and reproducible, although some cautionary notes are reported by (Hatfield et al., 1999). The digestion with acetyl bromide is performed using a heating block at 50°C for 2h as, according to (Hatfield et al., 1999), this combination of temperature and time is optimal for lignin solubilization, while minimizing saccharide release. To perform the quantification, 7mg of cell wall material prepared as above were weighed out in a 10mL Pyrex glass tube fitted with polypropylene cap. The walls of the tube were rinsed with 250µL of acetone to collect the cell wall material on the bottom of the tube, and acetone was evaporated overnight under a very gentle airflow.

To the sample 500µL of freshly prepared 25%(v/v) acetyl bromide solution in glacial acetic acid were gently added and tubes were capped and placed in a heating block at 50°C for 2h. After the two hours, samples were heated at the same temperature for an additional hour with vortexing every 15min. Subsequently, they were cooled on ice to room temperature. To each sample, 2000µL 2M NaOH and 350 µL of 0.5 M hydroxylamine hydrochloride (to ensure the decomposition of polybromide ions) were added, and samples were vortexed. Tubes were filled up exactly to the 10mL mark with glacial acetic acid, capped and inverted several times to mix. Then they were centrifuged to produce a particulate-free supernatant. 200µL of the supernatant was

pipetted into a 96-well plate, and their absorbance at 280nm was measured with a μQuant plate reader using KC4 software.

The amount of lignin in percentage was calculated from the absorbance using Equation 5.1 . Purified HCl-dioxane lignin from *Miscanthus* samples has a specific absorption coefficient (SAC) of 17.78 g^{-1} L cm^{-1} (Lygin et al., 2011).

Equation 5.1

$$ABSL(\%) = \frac{A_{280}^{0}}{SAC \cdot PL} \cdot \frac{V_R}{W_S} \cdot 10^{0}$$

Where:

ABSL(%) = percentage of acetyl bromide soluble lignin

A_{280} = Absorbance reading at 280nm

SAC = Specific Absorption Coefficient (17.78 g^{-1} L cm^{-1} for *Miscanthus*)

PL = pathlength determined for the 96-well microplates with a volume of 200μL per well used during the analysis (0.556cm) obtained from Costa (2015)

V_R = reaction volume in L

W_S = sample weight in g

5.2.7 Lignin composition

Lignin composition was determined according to (Robinson and Mansfield, 2009) as used by (Foster et al., 2010) with minor modifications from (José Carli, 2017). Approximately 2 mg of cell wall material was weighed out into a screw-capped glass tube for thioacidolysis. To each sample, 200 μL of Ethanethiol (EtSH), boron trifluoride (BF3), dioxane solution[5] were added. Vial headspace was purged with

[5] Prepare carefully the 2.5% boron trifluoride diethyl etherate (BF3), 10% ethanethiol (EtSH) solution. A balloon filled with nitrogen gas must be used to displace the lost volume in the dioxane bottle with nitrogen. Dioxane is very hazardous, do not take

nitrogen gas and capped immediately. Samples were heated at 100°C for 4 hours with gentle mixing every hour. The reaction was terminated by cooling on ice for 5 minutes. Subsequently, vials were decapped, and 150 µl of 0.4M sodium bicarbonate was added and vials vortexed. For the clean-up 1 ml of water and 0.5 ml of ethyl acetate were added, and after vortexing, left to stand for a minute to separate the different phases (ethyl acetate on top, water on bottom). For these intermediate steps vials were capped using a stripe of Parafilm. Afterwards, 150 µl of the ethyl acetate layer were transferred into a 200 µL vials and care was taken that no water was transferred. The ethyl acetate solvent was evaporated by leaving the tubes open overnight under the fume hood. The next day, 200 µl acetone were added and then evaporated (to remove excess water, this step was repeated twice). For the derivatization 30 µl of methoxyamine hydrochloryde dissolved in pyridine (20mg/mL) were added to the vials. After being crimp-capped, the vials were incubated for 15 min at 90°C in a heating block. The vials were then decapped and 20 µl of N,O-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) were added in order to derivatize the resulting compounds. The vials were recapped and incubated 15 min at 90°C.

Samples were analysed by gas chromatography/mass spectrometry on standard Agilent equipment comprising a 5973 network Mass Selective detector, a 6890 Series GC and a 7683 series autosampler. An Agilent HP-5MS column was installed (30 mm X 0.25 mm X 0.25 µm film thickness). The following temperature gradient was used

---

samples or equipment out of the hood. Volumes needed for the preparation of the solution per sample were: 175 µl dioxane; 20 µl EtSH; 5 µl BF_3.

with a 3 min solvent delay and a 1.1 ml/ min flow rate: Initial hold at 130 °C for 3 min; a 3°C/min ramp to a 250 °C and hold for 1 min; allow equilibration to the initial temperature of 130 °C. Peaks were identified by characteristic mass spectrum ions of 299 m/z, 269 m/z, and 239 m/z for S, G, and H monomers, respectively. The composition of the lignin components was quantified by setting the total peak area to 100%.

5.2.8 Saccharification efficiency

The amount of glucose, xylose and arabinose that can be enzymatically released from the cell wall was determined as described by (Resch et al., 2015) with some modifications. Approximately 10mg of cell wall material was weighed out in 2 mL micro-centrifuge tubes with screw cap[6] , and weight was recorded to the nearest 0.1mg. To each tube 300µL of 100% acetone were added to collect the CWM material to the bottom. Acetone was left to evaporate overnight. To each tube, 1 mL of saccharification mix was added. The saccharification mix was prepared in a way to contain 959 µL of de-ionised water, 30 µL of freshly prepared 1MCitrate Buffer (pH 5.2), 3 µL of Cellulase blend (Merck, SAE0020-50ML), 8µL of 1% Sodium Azide. Tubes were incubated for 72 hours in a shaking incubator at 150 rpm and 50 °C. When the digestion was completed, tubes were centrifuged at 10,000rpm - 5mins and supernatant collected in a new tube. Samples were then kept at -20 °C for a maximum of 3 days and monosaccharides were quantified as described in subsection 3.2.1.4.2.

[6] important to avoid evaporation during incubation

5.3 Results

5.3.1 Saccharification efficiency

Here we wanted to assess the variability in quality for the conversion of leaf and stem organs in 12 *Miscanthus* genotypes from different genetic backgrounds (9 natural accessions, 1 natural hybrid and 2 artificial hybrids).

To do so, the extracted cell wall material from leaf and stem was digested for 72 hours with a commercial enzymatic cocktail. The efficiency of saccharification of the biomass was measured by the amount of arabinose, glucose and xylose released in the supernatant. The values were the average of 4 technical replicates of the test. Results are presented in Table F. 1 as mean and standard deviation between the values of the three biological replicates. Significance of the differences between genotypes and plant portion was tested using ANOVA test and Tukey HSD as a posthoc test. Results of the analysis of variance are presented in Figure 5-1. More detailed results are presented in Table F. 1.

The effect of the interaction between genotype and plant portion on the amount of glucose and xylose enzymatically released was statistically significant (Table F. 9). Therefore, the organ-effect was tested separately on each genotype (Figure 5-1). There was a significant effect in glucose enzymatically released between leaf and stem portion in Bedivere, Gaheris, Galahad, Gareth, Kay, Lamorak, Lancelot and Tristan. In all the cases where the differences were significant, leaf samples were releasing more glucose than stem samples.

For the xylose release, there was a significant effect of the organ in Gaheris, Gawain, Kay, and Lamorak. In Gaheris, Kay and Lamorak the amount of xylose released from the leaf was higher than from the stem, while in Gawain, the amount was lower in the

leaf compared to the stem. In this regard, Gawain differentiates from the other genotypes of *M.floridulus* (Percival, Lancelot and Geraint) in this study.

The analysis of the saccharification efficiency performed separately for different organs allowed to detect genotype-specific differences hidden in the previous test on the pooled biomass from the same genotypes (section 3.3.2.1). These results, while confirming that the recalcitrance of the biomass to deconstruction is different between plant organs, they also show that when a broader range of species is considered, the direction and extent of the variability is affected by the specific genotype considered within the species.

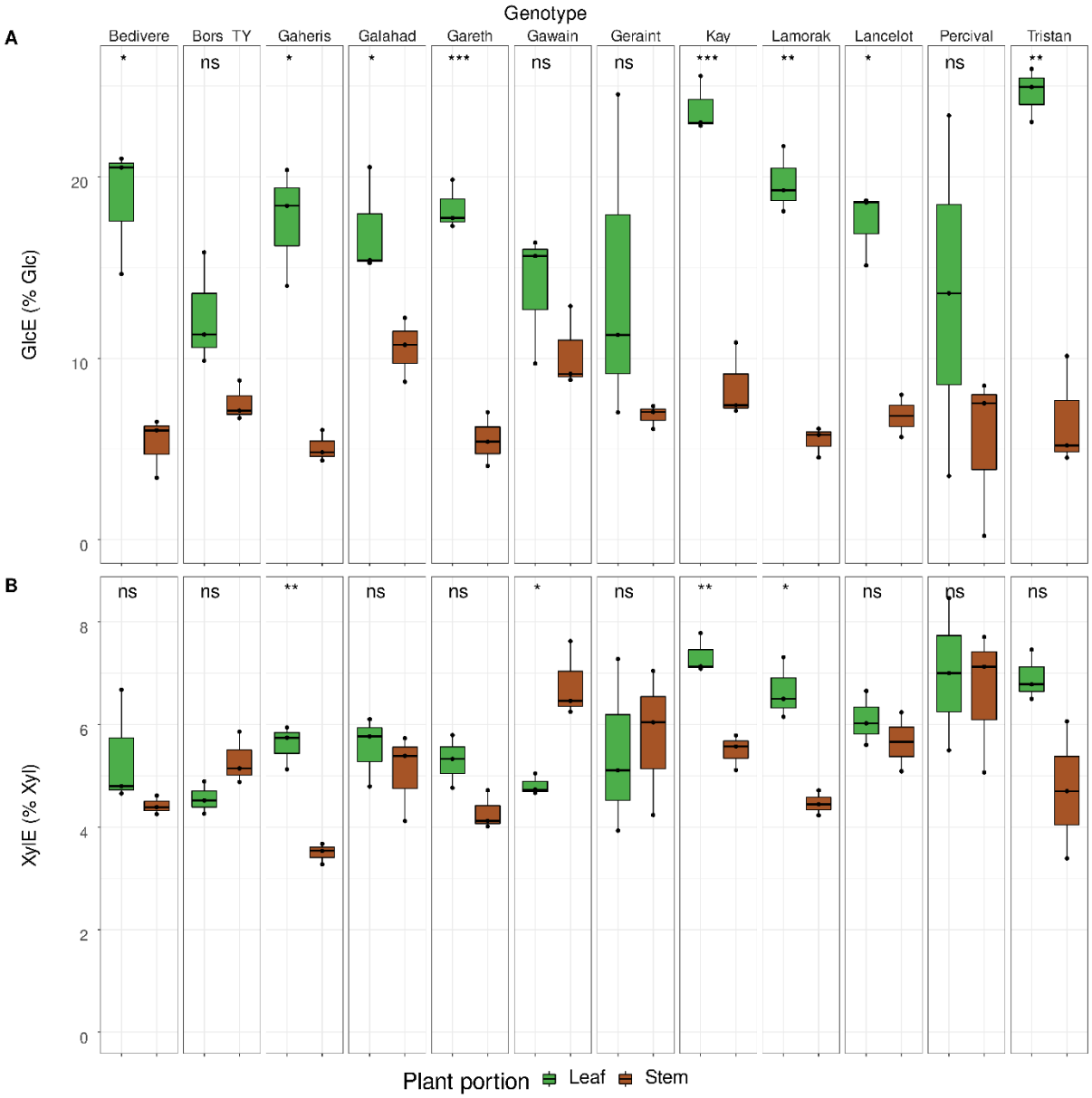


Figure 5-1 Effect of plant organ on the amount of each monosaccharide enzymatically released.
Sample of biomass form leaf and stem were digested for 72 hours with the CellTech enzymatic cocktail. Values are the average of four technical replicates on at least 3 biological replicates of each genotype in each condition. Values on the y-axis represent the amount of each monosaccharide expressed as a percentage in weight out of the total monosaccharide present in the cell wall. Boxes represent the mid two quartiles with the median drawn. The colour represents the plant portion (green = leaf, brown = stem) Leaf and stem of each genotype were compared using a t-test. Significance code ns: p > 0.05, *: p <= 0.05, **: p <= 0.01, ***: p <= 0.001,***: p <= 0.0001.

5.3.2 Matrix monosaccharides

The recalcitrance of the cell wall is a trait that can relate to the composition and structure of the cell wall. The hypothesis that differences between genotypes in the saccharification of the biomass from different plant portions were related to cell wall structure and composition was tested. Firstly, the composition of the hemicellulose was investigated. The matrix monosaccharides were extracted using a TFA method and quantified. In the supernatant, it was possible to detect fucose (Fuc), arabinose (Ara), glucose (Glc) galactose (Gal), xylose (Xyl) and mannose (Man). The values for the matrix monosaccharides were added together. The sum was used as the amount of hemicellulose in the cell wall (Hem). Values of the variates obtained on the four technical replicates of the test on each biological replicate were averaged, and values used to calculate the mean and standard deviation for the monosaccharide content in leaf and stem of each genotype. A summary of the results is presented in Table F. 2 and Table F. 3. The significance of the differences was tested globally using an ANOVA test and then with pairwise comparisons between leaf and stem of each genotype. Significant differences are represented in Figure 5-2, and more details are shown in Table F. 2 and Table F. 3.

The total amount of hemicellulose contained in the cell wall was not significantly different between genotypes and plant organs (Table F. 10).

There was no effect of the genotype on the amount of any of the matrix monosaccharides detected. Conversely, the effect of the plant portion was statistically significant for Ara and Gal. The amount of arabinose was 32% lower (Stem, mean = 2.252%, sd = 0.410; Leaf, mean = 3.330%, sd = 0.726, $p<0.005$) and galactose 36% lower (Stem, mean =0.70 , sd =0.182 ; Leaf, mean =1.108 , sd =0.240, $p<0.005$) in stem compared to leaves.

Considering that there were no differences in hemicellulose content, this result may indicate differences in the relative abundances of monosaccharides that make up hemicellulose.

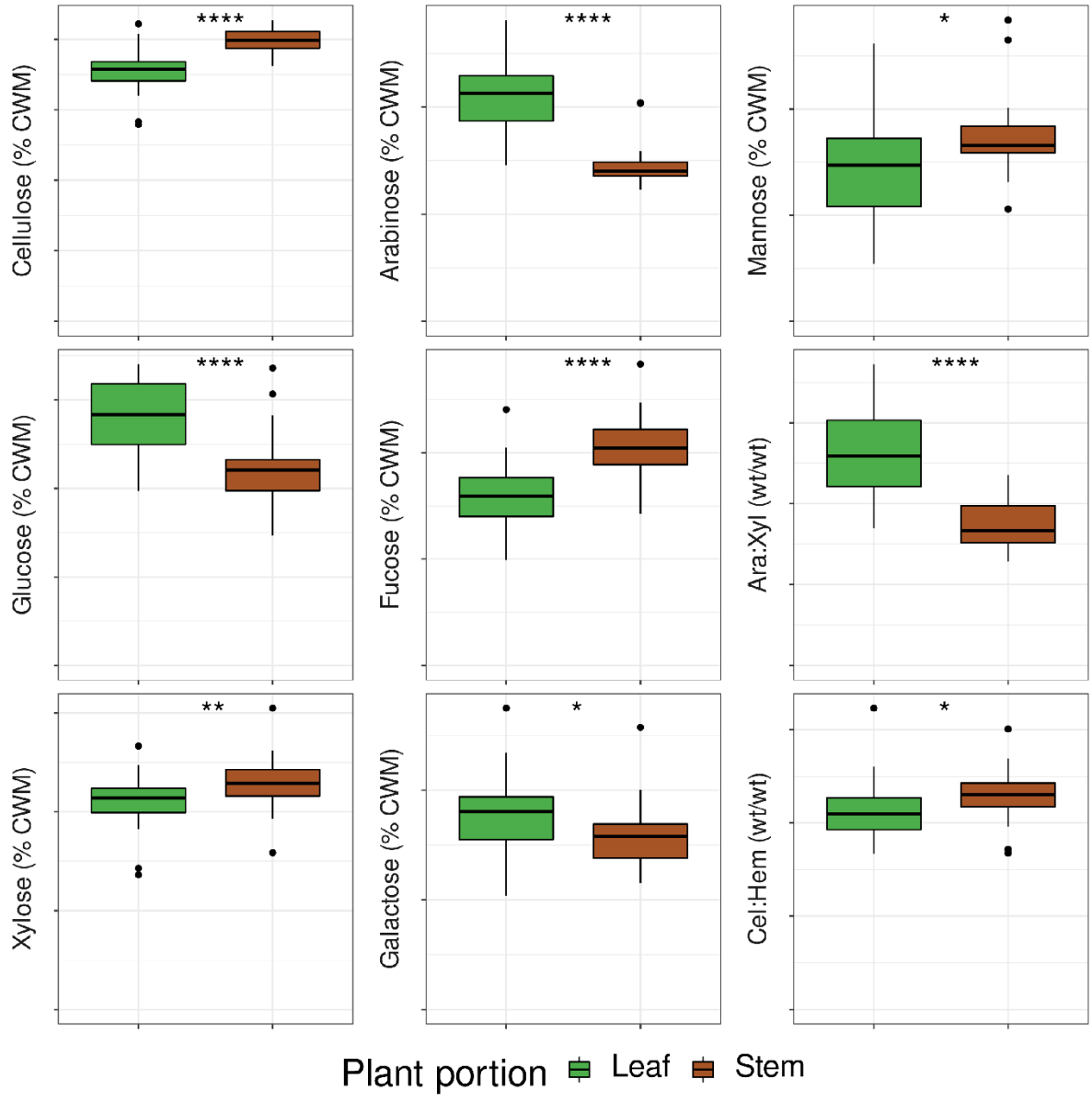


Figure 5-2 Effect of organ on matrix monosaccharides in the cell wall.
Matrix monosaccharides were quantified after hemicellulose was extracted using a treatment with TFA. There was a significant effect of genotype organ interaction and differences were tested independently for each genotype with a t-test. Values on the y-axis represent the amount of each monosaccharide expressed as a percentage in weight out of the total cell wall material. Boxes represent the mid two quartiles with the median drawn. The colour represents the plant portion (green = leaf, brown = stem) Significance code ns: p > 0.05, *: p <= 0.05, **: p <= 0.01, ***: p <= 0.001,***: p <= 0.0001.

5.3.3 Cellulose content

On the first attempt to quantify the crystalline cellulose, the pellet of the matrix used for hemicellulosic monosaccharide determination was purified and digested with single sulphuric acid treatment. However, this quantification failed. A significant amount of cellobiose was detected in the supernatant, indicating that the single acid digestion did not degrade the structure of the cellulose completely. Therefore, the amount of cellulose was estimated from the total amount of glucose released after double acid digestion of the CWM on a new set of samples. Values were averaged between technical replicates. Mean, and standard deviation for each plant portion on each genotype were calculated from the values obtained of the test repeated on the three biological replicates. Results are shown in Table F. 4. Differences were tested globally using a two-way ANOVA test and leaf, and stem of each genotype was compared using a t-test for significance. Results of the test are reported in Table F. 11 and presented inFigure 5-2.

The ANOVA test showed a significant effect of both plant portion considered and genotype on the cellulose content in the samples. However, the Tukey HSD test revealed that none of the genotypes was significantly different from the others. At the same time, the leaf had significantly less cellulose than the stem.

5.3.4 Cell wall architecture

Some proportions between monosaccharide contents in the cell wall were used as indicators of cell wall architecture. Here, the ratio between arabinose and xylose and between hemicellulose and cellulose were used. The degree of substitution of arabinose positively affects the enzymatic digestibility of cell wall (Li el al., 2013) and higher contents of cellulose in the cell wall have been associated with a reduced digestibility (da Costa et al., 2019). Therefore, differences in these two traits were

expected between the genotypes and organs showing different recalcitrance in this experiment. since they were reported as traits related to cell wall recalcitrance. A summary of the values obtained is presented in Table F. 5. Results of the statistical tests for differences are shown in Table F. 12. Statistically significant differences are presented in Figure 5-2.

Both the effect of genotype and plant portion had a significant effect on both parameters. However, the Tukey HSD test found that there was no difference between the genotypes. Leaf had a significantly higher Ara: Xyl and a lower Cel/Hem compared to the stem.

These results indicate that in stems, the cell wall contains a higher amount of cellulose, and the xylans chains have a lower degree of decoration with arabinosyl residues.

5.3.5 Lignin content and composition

Differences in recalcitrance have frequently been associated with the none-carbohydrate portion of the cell wall. Specifically, with the content and composition of the lignin component. The lignin in leaf and stem of the 12 genotypes in this study was extracted and quantified using the acetyl-bromide method. Its relative composition in terms of the three main monolignols was also determined. Both analyses were carried in 4 technical replicates and values were averaged. Mean, and standard deviation for each portion and each genotype was determined on the three biological replicates. Results are reported in Table F. 6and Table F. 7. Differences were globally tested using an ANOVA test (Table F. 13 and Table F. 14). Significant results are graphically presented in Figure 5-3, Figure 5-4 andFigure 5-5.

The amount of lignin was not significantly different between leaf and stem (Table F. 13), however, when differences were tested independently for each genotype, the lignin

content was significantly higher in stem samples compared with leaf samples for Gaheris and Percival (Figure 5-3).

Subsequently, the composition of the lignin from leaf and the stem of the 12 genotypes were investigated. The content of the monolignols guaiacyl (G) and p-hydroxyphenyl (H) was significantly lower in lignin from stem compared with the leaf. In contrast, syringyl (S) lignin was significantly lower in leaf compared with stems (Figure 5-4). In terms of genotype effect, for the content of G and S, only the comparison between Percival and Tristan showed a significant difference, with the latter having a higher percentage than the former for both monolignols (Figure 5-7). For the H component, the only significant comparison between genotypes was the one between Lamorak and Tristan.

Ratios between monolignols can help to infer the structure of the lignin polymers (Boerjan et al., 2013; Vanholme et al. 2010). The ratios H to G (H: G) and S to G (S: G) were considered. H: G was not significantly different between leaf and stem. At the same time, S: G was significantly higher in stem compared to leaf CWM (Figure 5-4). When genotypes were compared to each other with a Tukey HSD test, Lamorak was significantly different from Gawain, Lancelot, Tristan and Geraint. Geraint differed significantly from Bedivere, Bors, Gaheris, Galahad, Gareth, Kay and Percival.

Overall, these results show that lignin composition, in contrast with lignin content, was variable between plant portions. All the three components of the lignin considered separately were significantly different between leaf and stem, although in terms of ratios, only the ratio between S: G was.

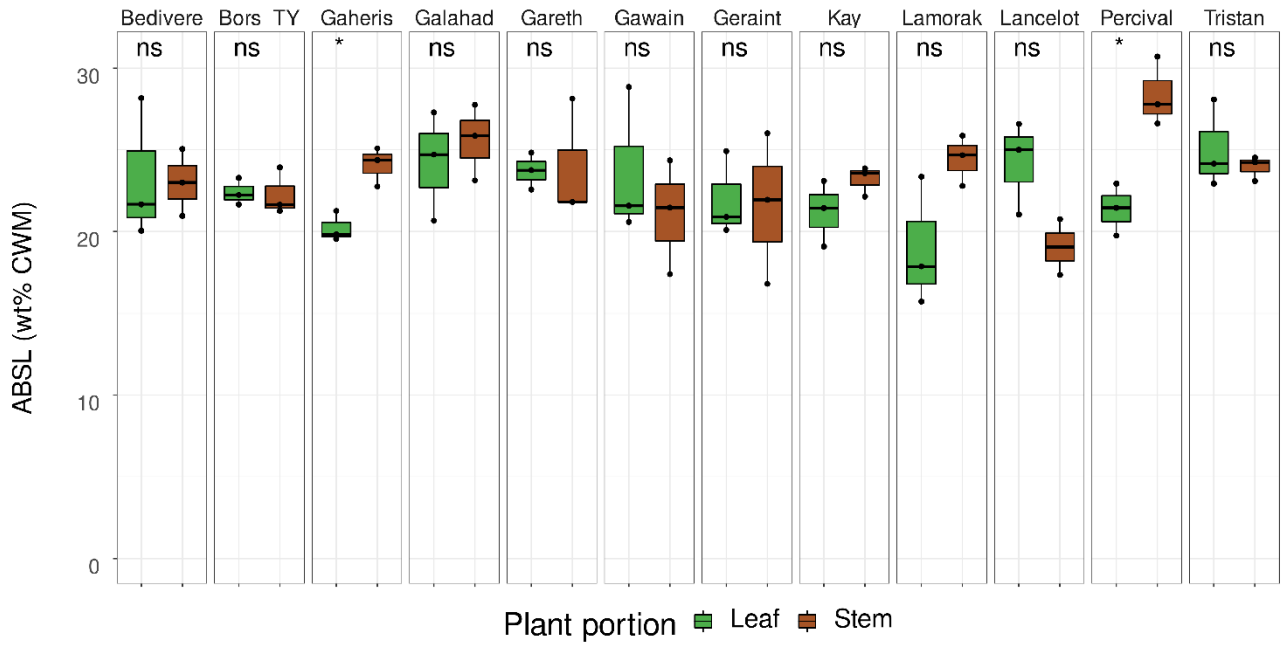


Figure 5-3 Effect of organ on the lignin content of the cell wall in the genotypes.
The total amount of lignin was quantified using the acetyl bromide extraction method. Values on the y-axis represent the amount of Acetyl bromide soluble lignin (ABSL) expressed as a percentage in weight out of the total cell wall material. The experiment was performed in 4 technical replicates. Values of the technical replicates were averaged for each of the three biological replicates. Boxes represent the mid two quartiles with the median drawn. The colour represents the plant portion (green = leaf, brown = stem) The interaction between genotype and organ was significant, and differences were tested independently for each genotype with a t-test. Significance code ns: p > 0.05, *: p <= 0.05, **: p <= 0.01, ***: p <= 0.001,***: p <= 0.0001.

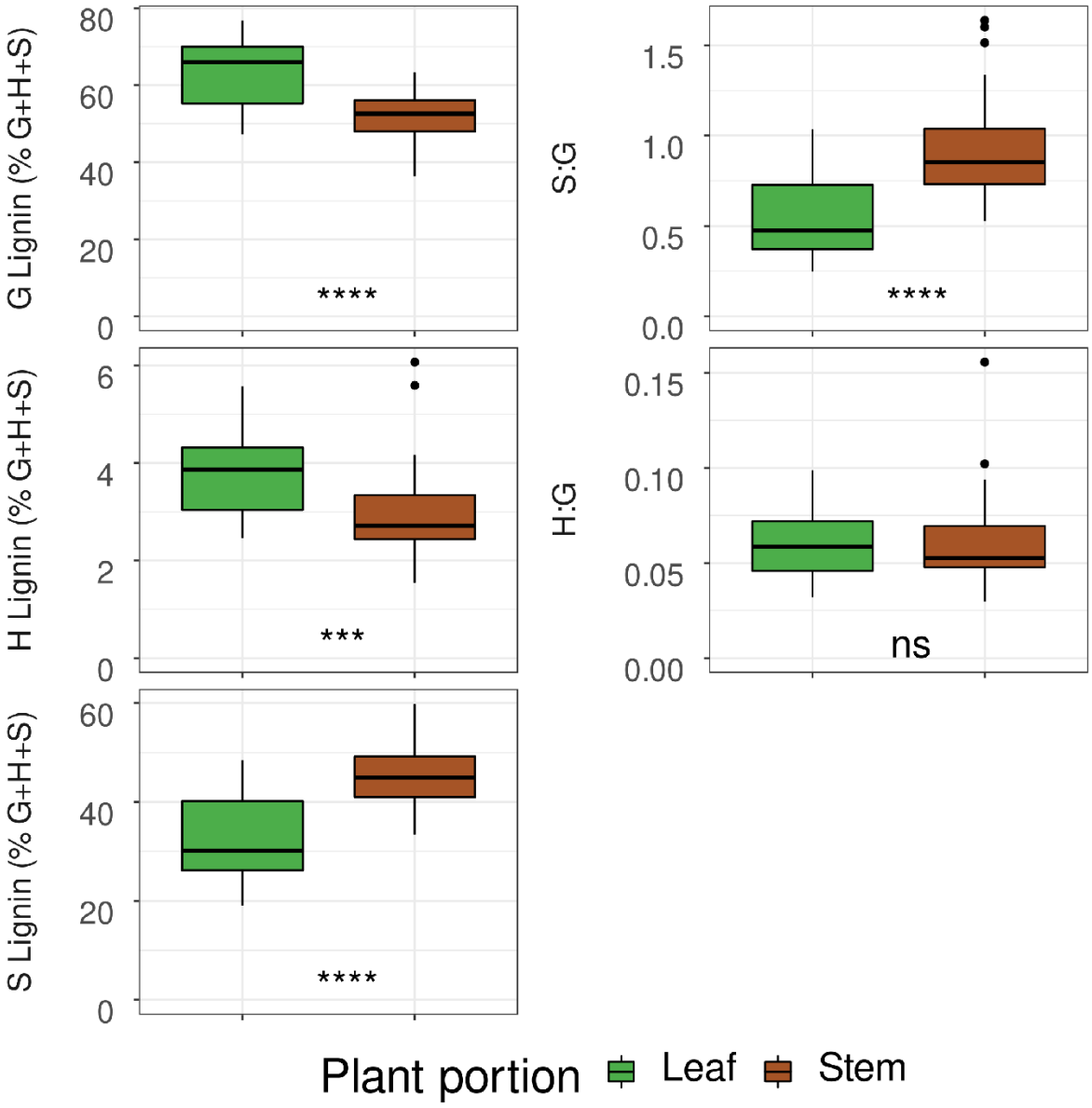


Figure 5-4 Effect of plant portion (leaf or stem) on the percentage of monolignols in the lignin of the genotypes.
The three phenylpropanoid units, p-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) of the lignin in the cell wall of leaf and stem portion of each genotype were quantified by GC-MS after lignin degradation by thioacidolysis. Each test was performed in four technical replicates on each biological replicate and values of the technical replicates were averaged for each of the three biological replicates (y-axis). Boxes represent the mid two quartiles with the median drawn. The colour of the box represents the plant portion (green = leaf, brown = stem). The points represent measurements of the biological replicates of the genotypes. Effect of plant portion was tested using a t-test. Significance is reported on the boxes. Significance code ns: p > 0.05, *: p <= 0.05, **: p <= 0.01, ***: p <= 0.001,***: p <= 0.0001.

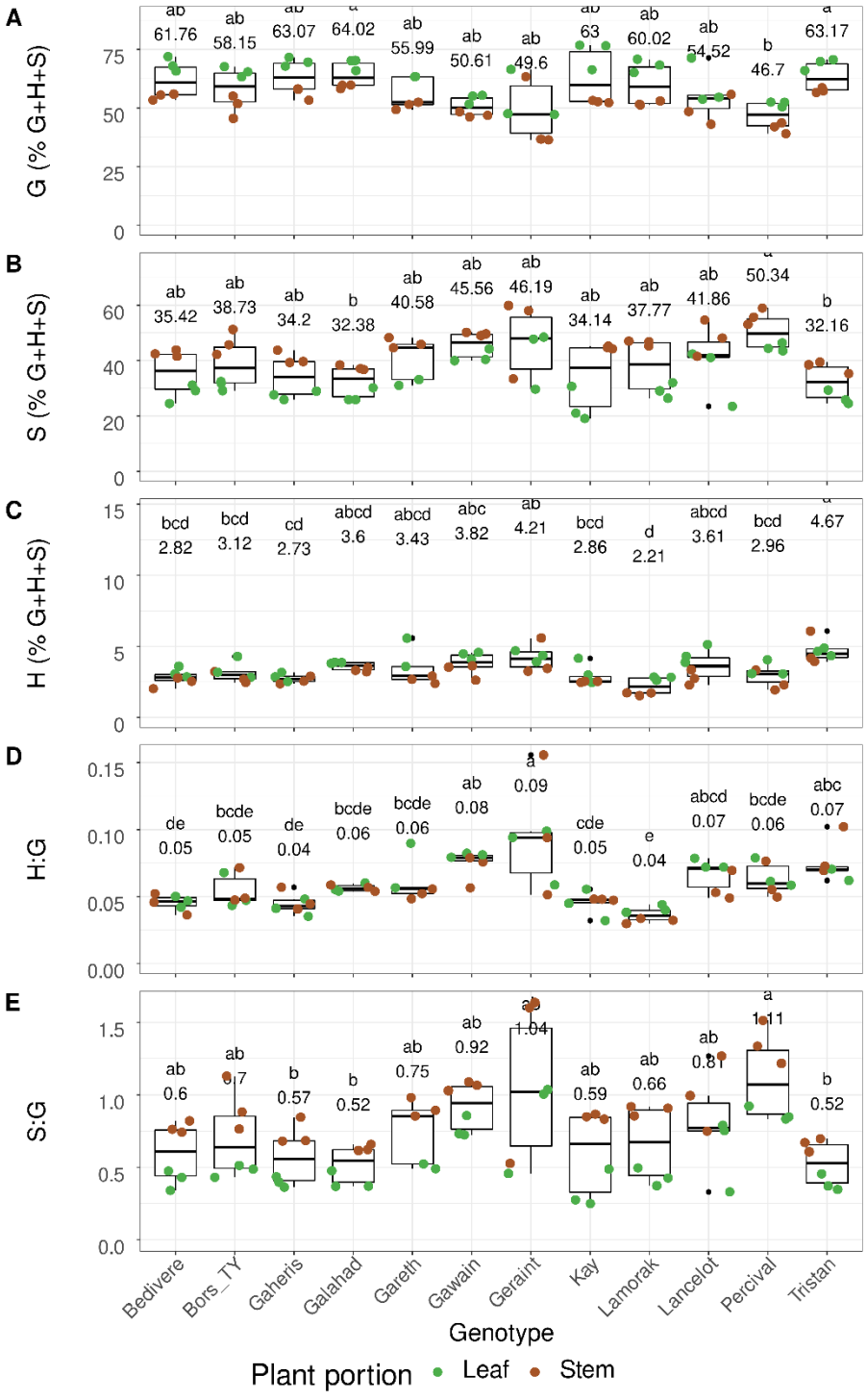


Figure 5-5 Effect of genotype on lignin composition.
The three phenylpropanoid units, p-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) of the lignin in the cell wall of leaf and stem portion of each genotype were quantified by GC-MS after lignin degradation by thioacidolysis. Each test was performed in four technical replicates on each biological replicate and values of the technical replicates were averaged for each of the three biological replicates (y-axis). Labels on the x-axes are the names of the genotypes in this experiment in alphabetical order. Boxes represent the mid two quartiles with the median drawn. The colours of the boxes represent the genetic group of each genotype. Points represent measurement on the biological replicates for leaf and stem (orange= stem; blue =leaf). Different letters symbolise significant differences according to the Tukey test of significance (α = 0.05).

5.4 Discussion

5.4.1 Saccharification efficiency

Traits connected directly with recalcitrance are challenging targets for breeding since alteration of the cell wall integrity can result in less resilient plants. Conversely, plant architecture can be more easily targeted to alter biomass quality for conversion without affecting plant biomass quality and production. The latter approach takes advantage of differences in biomass quality for conversion between leaf and stem and was proposed before (da Costa et al., 2017; da Costa et al., 2019). In this chapter, the differences in biomass quality were explored between two portions of the plant, the leaf and the stem, in 12 genotypes of the genus *Miscanthus*.

Previous experiments have used genotypes belonging to the three main species used in Europe, namely *M. x giganteus* and its parent *M. sinensis* and *M. sacchariflorus* (da Costa et al., 2017; da Costa et al., 2019; van der Weijde et al., 2017a; van der Weijde et al., 2017b). In the present study, *M. sacchariflorus*, *M. sinensis*, *M. floridulus*, *M. lutarioriparius* and *M.* x *giganteus*, natural and artificial were used, expanding the knowledge previously reported in the literature.

Firstly, the saccharification efficiency was determined in samples of CWM from leaf and stem. Using the amount of ethanol produced, da Costa et al. (2017), using plants of *M. x giganteus, M. sinensis* and *M. sacchariflorus* did not find differences in the recalcitrance between leaf and stem senesced material. Mangold et al. (2019) compared the amount of methane per hectare that could be produced by genotypes of *M. sinensis*, *M. x giganteus* and artificial hybrids between *M. sinensis and M. sacchariflorus*. They found that differences in the gas produced from leaf and stem were related to the genotypes considered. In both the studies above, there was no

quantification of the monosaccharides that can be enzymatically released from the cell wall of *Miscanthus*. Le Ngoc Huyen et al. (2010) using different enzymatic cocktails, quantified cellulose and arabinoxylan. Using the amount of hexoses and pentoses released, they found that in *M.* x *giganteus* leaves and sheaths were 2.5 times less recalcitrant to deconstruction compared to the stem internodes.

The results in the present study corroborated the difference between leaf and stem observed by other authors using a broader range of species of *Miscanthus*.

The analysis performed separately for different organs allowed to detect differences in cell wall characteristics between genotypes with different recalcitrance not detected in the previous test on the pooled biomass (section 3.3.2.1).

This knowledge could impact on future breeding programmes aimed at improving *Miscanthus* biomass quality for conversion. It is possible to imagine including the leaf to stem ratio in quality indexes in progeny screening or parental selection. When using genetic breeding, organ-specific promoters of genes involved in biomass deconstruction could be considered for investigation.

5.4.2 Matrix monosaccharides

Next, the cell wall composition in leaf and stem was tested, to understand if the source of the different conversion qualities observed could be related to specific differences in cell wall composition and structure. Commelinid monocots contain a cell wall matrix composed mainly of arabinoxylans (AXs), xyloglucans (XGs), and mixed-linkage (1→3),(1→4)-β-d-glucans (MLGs) (Buchanan et al., 2000).

In the present study, while the hemicellulose in the cell wall did not differ significantly between leaf and stem in amount, it was differing in composition.

The differences resided in the amount of arabinose and galactose (See Fig. 5.2) that were lower in stem compared to leaf material. Arabinose, in the grass cell wall, is

mostly deriving from AXs (Buchanan et al., 2000). Differences could be related to differences in tissue distribution between organs since L-arabinans have been observed to be abundant in the stomata companion cells (Cao et al., 2014).

The same difference observed here for galactose was observed in *M. giganteus*, *M. sinensis* and *M. sacchariflorus* previously (Le Ngoc Huyen et al., 2010). Galactose in the cell wall of *Miscanthus* derives from homogalacturonan. The abundance of homogalacturonan epitopes was higher close to vascular bundles and related to the water and nutrient transport efficiency (Cao et al., 2014). Therefore, the variability in the extent of difference of galactose between leaf and stem material in this experiment is most likely related to different tissue organisation between leaf and stem organs. More information, in this case, could derive from the localization of the monosaccharides in the intact cell wall of the genotypes.

Information about the composition of the cell wall matrix could inform future breeding programs aimed at the production of new varieties of biomass crops producing lignocellulosic biomass for specific conversion processes.

5.4.3 Cellulose

Looking at the cellulose content, it was seen that the amount of cellulose in the cell wall was higher in stems compared to leaf (Section 5.3.3). This result is not surprising since a similar difference has been reported before and was suggested to be associated with the load-bearing function of the stem (da Costa et al., 2014). The paradox for the conversion is that cellulose is the primary source of glucose. Still, its crystallinity increases the recalcitrance of the cell wall to deconstruction (Zoghlami et al., 2019).

5.4.4 Cell wall architecture

The structural organisation of polysaccharides in the cell wall plays a pivotal role in determining the recalcitrance to deconstruction (Li et al., 2013; da Costa et al., 2014;

van der Weijde et al., 2017b; Da Costa et al., 2019). Some proportions between monosaccharide amounts can be used to infer about the cell wall organisation. The ratio between arabinose and xylose is an indicator of xylan arabinose decoration. It has been related to cell wall recalcitrance (Li et al., 2013). The result indicates that in stems, the cell wall contains a higher amount of cellulose, and the xylan chains have a lower degree of decoration with arabinosyl residues. It has been suggested that the level of decoration and the pattern of decoration of xylans can play a crucial role in the interaction between hemicelluloses and cellulose. It has been proposed that the decoration of xylans can determine the extent to which the AXs chains can interlink with cellulose fibrils connect them to lignin and matrix (Cass et al., 2016). Therefore, the observed lower level of xylan decoration (Section 5.3.4) and higher content of cellulose in the stems (Section 5.3.3) is in agreement with the higher recalcitrance of this portion.

5.4.5 Lignin content and composition

Lignin can increase the biomass recalcitrance and has been often associated with a lower quality of the biomass for biochemical conversion. However, it has been shown that recalcitrance is a complex trait. Although lignin content contributes, it is not the only determinant (Da Costa et al., 2019). In the present study, this was confirmed. There was a limited variation in the content of lignin between plant organs and genotypes, which suggests that other factors had a significant role in determining the effect of organ on cell wall recalcitrance.

Lignification has different roles in the plants. It is necessary to provide the plat structure with support to stand, but also protects the tissues from external biotic and abiotic agents. In contrast with the low variability of lignin content, the lignin composition, in terms of monolignols, was varying between leaf and stem. All the three

components of the lignin, considered separately, were significantly different between leaf and stem. This could be a consequence of the different role of the two organs in the plant architecture and the fact that lignin with different composition is required for the main function (protection in the leaves, support in the stems).

In terms of ratios between the monolignols, only that between S: G was significantly different. Recently, Bhatia et al. (2019) expressed the ZmMYB167 gene, encoding a transcription factor, under the control of the constitutive maize ubiquitin promoter (ZmUbi1) in *Brachypodium distachyon*. ZmMYB167 is a potential activator of phenylpropanoid biosynthesis. Its expression resulted in a higher content of syringyl (S) lignin monomers (~11% to 16%), lower levels of guaiacyl (G) lignin monomers (~17% to 25%), and a concomitant increase in the S/G ratio (~32% to 53%) compared with the wild type. Biomass from the mutants released 20% less glucose when digested for 72 hours enzymatically. Similarly, it cannot be excluded that differences in the level of expression of genes in the pathway of phenylpropanoid production could be related to the differences observed in the present study. Future investigation should try to address the extent of intra- and interspecific variability of gene expression in the *Miscanthus* genus and the relation with stress perception and adaptation.

6 Investigating the effect of nutrient stress on *Miscanthus* spp.

6.1 Introduction

Nutrient limitation is receiving increased attention for the opportunity of growing biomass crops without interfering with food production because it is one of the abiotic interactors that can limit plant growth and productivity on marginal lands. Recent research indicates that mixtures of perennial feedstocks represent a suitable cropping system for biomass production on marginal lands (Carlsson et al., 2017; Mehmood et al., 2017; Robertson et al., 2017). To profitably exploit marginal areas for the production of biomass of adequate yield and quality the stability of these traits under variable abiotic conditions must be improved. Considering the importance of cell wall characteristics in determining the quality of the biomass, it is crucial to investigate the effect of abiotic stress on this cell component. The effect of cold, salinity and lodging drought stresses and heavy metals on cell wall characteristics is well documented (Le Gall et al., 2015; van der Weijde et al., 2016; Pancaldi and Trindade, 2020). More limited is instead, the information about the effect of nutrient stress.

The literature on this aspect has been recently reviewed by Ogden et al. (2018). The authors observed that the cell wall determines the morphology of cells and tissues and nutrient availability drives changes in plant growth. Consequently, the cell wall structure should be directly or indirectly controlled by the nutrients available in the plant. It is known that most of the transport of nutrients from the soil to the apoplast involves their passage through Casparian strips, bands of highly lignified, hydrophobic cell wall material that encompass the radial and transverse domains of endodermal

cells in the zone of maturation of the root (Geldner, 2013). Thus plants regulate nutrient uptake through modulation of the cell wall permeability (Ogden et al., 2018). Recently da Costa et al., 2018 reported for the first time the effect of nutrient stress on cell wall recalcitrance on three *Miscanthus* genotypes. After 56 days of growth under nutrient stress conditions, biomass from *M.* x *giganteus* and *M. sacchariflorus* had reduced recalcitrance to enzymatic deconstruction, while the one from *M. sinensis* presented increased recalcitrance. Moreover, the differences were not related to matrix polysaccharides, or in cellulose and lignin contents. The authors attributed this result to the effects of nutrient stress on the fine structure of cell wall polysaccharides. This evidence and the lack of other studies pointed to the need for investigating the effect of nutrient stress in more detail. To better understand the physiological bases of the effects of nutrient stress on cell wall recalcitrance in *Miscanthus*, it was decided to induce single nutrient stress on the set of *Miscanthus* genotypes selected in chapter 3. This collection includes genotypes with contrasting characteristics in terms of the cell wall composition. The group also included three hybrids, which allowed to test the response of hybrids to abiotic stress conditions.

The investigation was focused on nitrogen and phosphorous limitation, for several reasons connected to their role in the sustainability and importance in the metabolism of the plant. Nitrogen is the second most abundant element in plants after carbon. It is found in proteins, nucleic acids, chlorophyll, co-enzymes, phytohormones and secondary metabolites (Hawkesford et al., 2011). Plants take up nitrogen from the soil as nitrate (NO_3) and ammonium (NH_4). Nitrogen limitation has been reported to affect the cell wall (Landi and Esposito, 2017). However, the molecular mechanisms by which N status affects cell wall biosynthesis and organization are still the object of study (Ogden et al., 2018).

Phosphorus has crucial importance in plant metabolism. It is a component of critical macromolecular structures (e.g. nucleic acids), involved in mechanisms of energy transfer (e.g., ATP) and regulating enzyme reactions via compartmentalisation.

The typical effect of phosphorous limitation in plants is an alteration of root architecture and leaf expansion. In particular, a concomitant reduction of the primary root and increase in the number and length of lateral roots has been reported. On a cellular level, phosphorus limitation can affect both cell production and cell division rate (Assuero et al., 2004). A detailed transcriptomic analysis revealed that phosphorus limitation was associated with genes controlling pectin modification, cell wall relaxation, hemicellulose/cellulose modification, and carbohydrate hydrolytic enzymes (Hoehenwarter et al., 2016). It has been hypothesised that phosphorus limitation affects root morphology through a cascade of reactions that involve the accumulation of iron in the apoplast that in turn affects cellulose, pectins and callose synthesis (Müller et al., 2015).

Phosphorous importance as a nutritive element for plants is a consequence not only of its functions in the cell metabolism, but also its relatively low abundance in the environment (Elser, 2012). Phosphorous as fertilizer is a non-renewable resource obtained from–P-rich geological deposits localized in just a few countries around the world. A phosphorous peak, similar to the concept of oil peak, has been hypothesized to happen in 2035 (Cordell et al., 2009). Developing crops and agriculture management systems that require less P while maintaining productivity will be crucial for plant production when reserves will be depleted (Gilbert, 2009), including biomass crops.

Here the results of an experiment aimed to understand the effect of nitrogen and phosphorus limitation on the carbohydrate composition of the cell wall of 12 genotypes of *Miscanthus* are presented.

6.2 Material and methods

6.2.1 Plant growth

6.2.1.1 Rhizome harvesting

Rhizomes of the selected 12 *Miscanthus* genotypes were harvested from ABR33 in January 2018. The field trial was described in chapter 2. For each genotype, a quarter of the available rhizomes of each plant in the three biological replicates was collected. Straight after collection rhizomes were cleaned with tap water to remove any residual of soil and stored at 4°C in moist clean sand similarly to the procedure reported by other authors (Ings et al., 2013).

6.2.1.2 Rhizome preparation

In April 2018, rhizomes were weighed and cut into portions of similar weight before being planted for germination. For each rhizome pictures of the full rhizome and of the sub-portions were collected along with total weight and weight of the single portion. For each genotype, one rhizome portion was stored to -80°C for mineral nutrient determination. To reduce the effect of rhizome size on the impact of the treatment, the partitioning of the rhizome pieces was carried out in a way to have similar weights for the individual pieces. Rhizome pieces from the same genotype coming from different biological replicates were kept separate in this phase to facilitate the observation of differences related to the site of collection. The pictures collected are in subsection F3.

6.2.1.3 Rhizome germination

Rhizomes were planted in a bed of 3cm of clean 100% vermiculite. They were then covered with another 2cm of the vermiculite. The substrate was maintained moist but not wet until germination, and the humidity was assessed by eye daily. Containers with the rhizomes were placed in a glasshouse with a 16h photoperiod and without temperature control. However, temperature and humidity were monitored using a Tinytag (Gemini Data Loggers) (Figure 6-1).

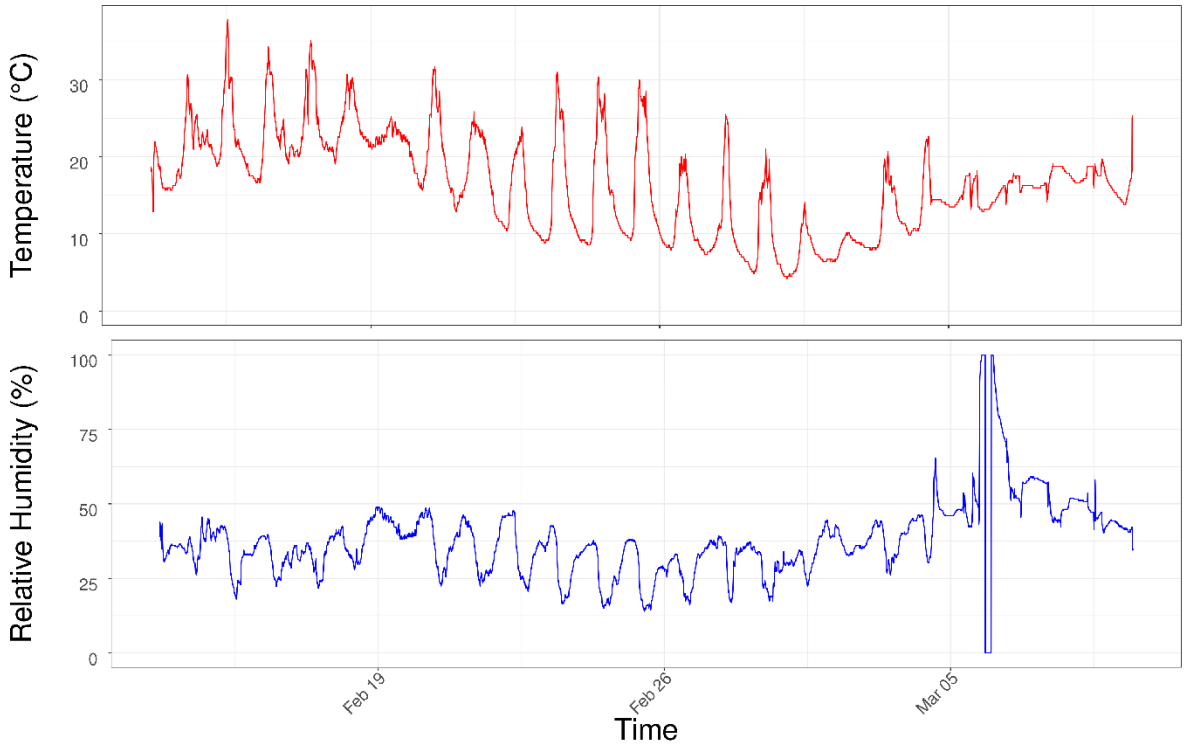


Figure 6-1 Line plot of temperature and humidity conditions during the germination
Temperature and relative humidity were monitored in the greenhouse using a MiniTag device with 15 minutesintervals during the germination period.

6.2.1.4 Experimental setup

The experimental design was completely randomised with five biological replicates of each treatment for each genotype. The replication was performed by assigning 5 plantlets of each genotype to each treatment. After germination, healthy plantlets and with similar size for each genotype were selected.

6.2.1.5 Planting

Germinated rhizomes were planted in 10” pots with saucers filled with ultralow nutrient compost with 10% vermiculite and 30% sand. The details of the characteristics of the ultra-low compost can be found in Table 6-1. Using a low nutrient compost ensures that while none of the structural characteristics of the soil is limiting for the plant, the amount of nutrient supplied can be finely controlled. For each pot, the total weight of the soil (including pot weight) and the weight of the plantlet were recorded. For the whole duration of the treatment and until harvest, the photoperiod in the glasshouse was maintained at 16hours of light and 8 hours of dark. Temperature and humidity were not controlled in the greenhouse. Environmental conditions, including measurements of humidity and temperature every 15 minutes, were monitored for the whole duration of the treatment and until harvesting using a TinyTag datalogger (Gemini Dataloggers) (Figure 6-2).

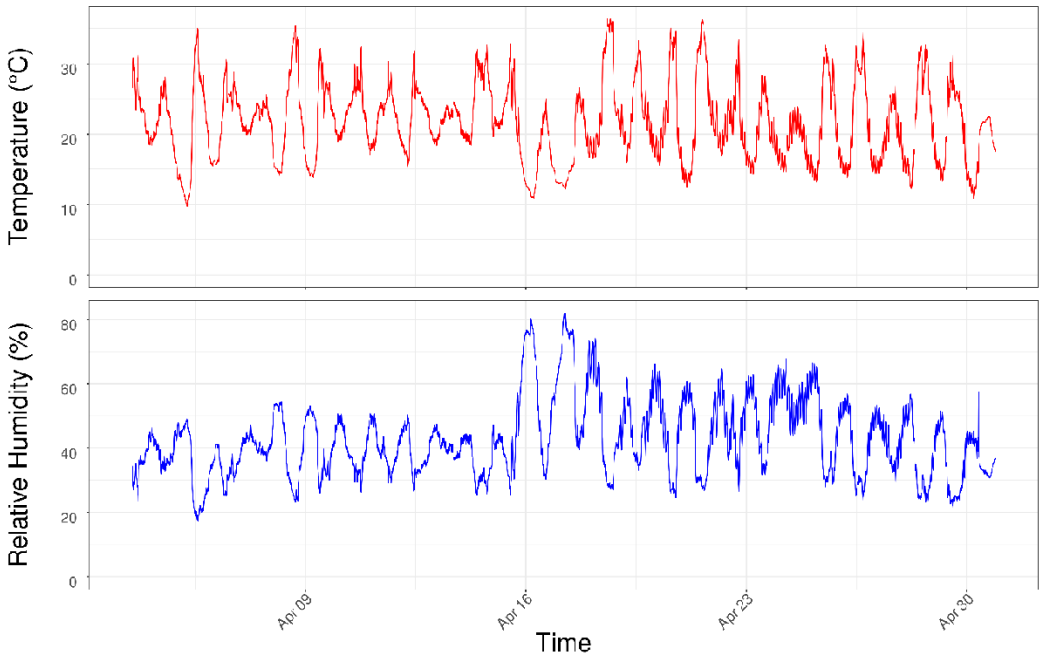


Figure 6-2 Line plot of temperature and humidity conditions during the treatment
Temperature and relative humidity were monitored in the greenhouse using a MiniTag device with 15 minutesintervals during the treatment.

Table 6-1 Characteristics of the low nutrient compost.

Details of the chemical composition of the low-nutrient compost used in mix with vermiculite to grow the plants in this experiment.

Param.	Unit	Value
Dens.	g/l	264
pH		3.6
Cond.	µs/cm	16
NH_4-N	mg/l	6.2
NO_3-N	mg/l	0.8
N (Tot.)	mg/l	7
Cl	mg/l	10.5
K	mg/l	1.2
Mg	mg/l	<0.2
Ca	mg/l	0.7
Na	mg/l	5.3
Fe	mg/l	0.13
P	mg/l	<1
Cu	mg/l	<0.01
Mn	mg/l	<0.01
Zn	mg/l	0.03
B	mg/l	<0.05
SO_4	mg/l	8.8

6.2.1.6 Nutrient solution

The substrate used for the growth was ultra-low in all the nutrients (Table 6-1). For the duration of the experiment, nutrients were supplied with fertigation once a week with a complete nutrient solution. The nutrient solution was prepared according to

(Hoagland et al., 1950)with the modification that allowed to fit the amount of each nutrient necessary in the amount of water provided during the irrigation. The amount of water provided with each irrigation was calculated taking in account the field capacity of the compost. The water content in the soil was maintained above the 75% of field capacity. The amount of each nutrient provided during the duration of the experiment was calculated in a way that it was equivalent to the amount of nutrient that the plant would have had if it was growing in the same amount of nutrient-rich compost. Since it has been observed that *Miscanthus* can grow well for a whole year in John Innes 3 compost, this compost was used as a reference for the calculation. The amount of compounds used for the preparation of the stock solutions and the nutrient solution for the control treatment is reported in Table 6-2, while the composition of the same solution without sources of nitrogen and phosphorus is reported in Table 6-3 and Table 6-4 respectively. The nutrient content of the final solutions is shown in Table 6-5

Table 6-2 Mineral composition of the complete double-strength Hoagland solution.

Full Hoagland							
Salt	**MW**	**Sol**	**Stock**	**Stock**	**Amount**	**Tot**	**Final**
	g	**g/L**	**mol/L**	**g/L**	**mL/1L**	**mL/30L**	**mM**
NH_4NO_3	80.0434	1500	1	80.0434	10	300	5
NaH_2PO_4	141.96		0.5	70.98	7	210	1.75
KCl	74.5513	316	1	74.5513	3	90	1.5
$MgSO_4.7H_2O$	246.4746	100	1	246.4746	4	120	2
$CaCl_2*6H_2O$	219.0757	2	0.5	109.5378	4	120	1.5
$MnSO_4.4H_2O$	223.0618		0.004483	1			0.045
	Microelements						
$ZnSO_4.7H_2O$	287.5496		0.000696	0.2			0.0007
H_2MoO_4 (molybdic acid)	161.9735		0.000617	0.1			0.0006
$CuSO_4.5H_2O$	249.685		0.000401	0.1	2	60	0.0004
H_3BO_3	61.833		0.008086	0.5			0.0081
							0.0001
KI	166.0028		0.000181	0.03			8
$NiSO_4$	154.756		0.000646	0.1			0.0001
EDTA-Fe-Na	367.05		0.1	36.705	2	60	0.1

Table 6-3 Mineral composition of the double-strength Hoagland solution lacking of sources of nitrogen.

	Without nitrogen						
Salt	**MW**	**Sol**	**Stock**	**Stock**	**Amount**	**Tot**	**Final**
	g	**g/L at 20 °C**	**M (mol/L)**	**g/L**	**mL/1L**	**mL/30 L**	**mM**
NH_4NO_3	80.0434	1500	1	80.0434	0	0	0
NaH_2PO_4	141.96		0.5	141.96	7	210	1.75
KCl	74.5513	316	1	74.5513	3	90	1.5
	246.474			246.474			
$MgSO_4.7H_2O$	6	100	1	6	4	120	2
	219.075			109.537			
$CaCl_2*6H_2O$	7	2	0.5	8	6	180	1.5
	223.061		0.00448				
$MnSO_4.4H_2O$	8		3	1			0.045
Microelements							
	287.549		0.00069				
$ZnSO_4.7H_2O$	6		6	0.2			0.0007
H_2MoO_4 (molybdic			0.00061				
acid)	161.9735		7	0.1			0.0006
			0.00040				
$CuSO_4.5H_2O$	249.685		1	0.1	2	60	0.0004
			0.00808				
H_3BO_3	61.833		6	0.5			0.0081
	166.002						0.0001
KI	8		0.000181	0.03			8
			0.00064				
$NiSO_4$	154.756		6	0.1			0.0001
EDTA-Fe-Na	367.05		0.1	0.1	2	60	0.1

Table 6-4 Mineral composition of the double-strength Hoagland solution lacking of sources of phosphorus.

	Without phosphorus						
Salt	MW	Sol	Stock	Stock	Amount	Tot	Final
	g	g/L at 20 °C	M (mol/L)	g/L	mL/1L	mL/30 L	mM
NH_4NO_3	80.0434	1500	1	80.0434	10	300	5
NaH_2PO_4	141.96		0.5	141.96	0	0	0
KCl	74.5513	316	1	74.5513	3	90	1.5
	246.474			246.474			
$MgSO_4.7H_2O$	6	100	1	6	4	120	2
$CaCl_2*6H_2O$	219.0757	2	0.5	109.5378	6	180	1.5
	223.061						
$MnSO_4.4H_2O$	8		0.004483	1			0.045
			0.00069				
Microelements			6	0.2			0.0007
	287.549						
$ZnSO_4.7H_2O$	6		0.000617	0.1			0.0006
H_2MoO_4							
(molybdic acid)	161.9735		0.000401	0.1	2	60	0.0004
			0.00808				
$CuSO_4.5H_2O$	249.685		6	0.5			0.0081
							0.0001
H_3BO_3	61.833		0.000181	0.03			8
	166.002		0.00064				
KI	8		6	0.1			0.0001
$NiSO_4$	154.756		0.1	0.1	2	60	0.1
EDTA-Fe-Na	367.05		0.1	0.1	2	60	0.1

Table 6-5 Mineral content of the nutrient solution in the three treatments

Element	Full	No N	No P
Unit	ppm	ppm	ppm
N	0.14007	0	140.07
P	0.0542045	54.2045	0
K	0.058647	58.647	58.647
Ca	0.040078	60.117	60.117
Mg	0.04861	48.61	48.61
S	0.06413	64.13	64.13
Mn	2.47221	2.47221	2.47221
Zn	0.045766	0.045766	0.045766
Cu	0.0254184	0.0254184	0.025418
Mo	0.057576	0.057576	0.057576

6.2.1.7 Field capacity determination

The field capacity of the compost mixture used for the experiment was measured as described by (Manmathan and Lapitan, 2013). One Whatman #2 filter paper was placed on the bottom of each of three plastic pots (9 x9 x 9.5cm, TEKU ref. 0350). The mass of the pot with filter paper was determined to the nearest 0.01g and recorded. Each pot was filled gently with 0.460 mL of compost mixture, and the weight of the pot and the soil was determined to the nearest 0.01g and recorded. The pots were placed into a shallow pan of water, allowing only the bottom few centimetres to become wet. The pots were allowed to become saturated from the bottom to the surface. Pots were then allowed to drain by gravity, and the weight was measured and recorded after 3 hours. Pots were allowed to dry completely over a period of 12 days. These pots were weighed again. Finally, the dry weight of the soil was determined by placing the pots in an oven at 105°C until constant weight.

Protocols reported in literature usually start with oven-dried compost. However, a first attempt of quantifying field capacity was unsuccessful because of the inability of the oven-dried compost of adsorbing water, maybe as a consequence of soil structure degradation caused by the drying procedure

The dry weight soil moisture fraction at the field capacity point was calculated according to (Walker, 1989)as follows:

Equation 6.2

$$W_{fc} = \frac{WW_{fc} - DW}{DW}$$

where:

W_{fc} = Weight of soil moisture fraction at the field capacity

WW_{fc}= Weight of the soil at the field capacity

DW= Dry weight of the soil

The volumetric field capacity was then calculated according to Equation 6.3. In the equation the weight soil moisture fraction was converted into volumetric moisture content by multiplying by the bulk density and dividing by the specific weight of water

Equation 6.3

$$q_{fc} = \frac{g_b \cdot W_{fc}}{g_]}$$

where:

q_{fc} = Volumetric field capacity

g_b= bulk density of the soil calculated as in Equation 6 .4

W_{fc} = dry weight moisture fraction at the field capacity point as calculated with Equation 6 .2

$g_]$= density of the water, assumed to be 1

Equation 6.4

$$g_{b=\frac{W_b}{V}}$$

where:

W_b= the dry weight of the soil

V= the volume of soil used

The field capacity of the compost mixture calculated in this way was used for determining the volume of water solution for the fertigation.

The field capacity measures the highest limit of available water. It is the moisture of the soil after drainage of the water contained in the macropores by gravity action. The measurement showed that the field capacity of the compost was approximately 63% of the compost weight. Consequently, the watering level was maintained above 47.25% of the compost weight (63 x 0.75). The average weight of soil in each pot was 3,000 g. This means that the compost was able to store ~1,500 g of water. Each fertigation was

performed using 1,000 mL of nutrient solution containing a fraction of the total amount of each nutrient that was meant to be provided during the 8 weeks of experiment.

Table 6-6 Field capacity determination in the compost.
Wc is the weight of the container, Wi is the initial total weight of the pot and soil, Wf is the final weight. The field capacity (F_c) was calculated as percentage of weight variation over the initial weight.

Wc	Wi	Wf	F_c
9.8	148.95	238.51	60.12756
10.06	148.07	239.35	61.64652
9.87	153.52	250.45	63.13835
Average			61.63748

6.2.2 Plant measurements

6.2.2.1 Phenological measurements

Phenological measurement included the determination of the length of the plant, the length and the width of the youngest completely expanded leaf. Measurements were taken according to Ings et al. (2013). Stem length was measured twice a week from soil level to the highest fully expanded leaf using a graduated ruler. The rate of elongation was then calculated using these measurements. Leaf length was measured on leaf 0 with leaf length (from the ligule to leaf tip) and width (midway between ligule and tip) measured using a graduated ruler. Leaf area (LA) was calculated as described (Clifton-Brown and Lewandowski, 2000):

$$LA = l_c \cdot LL \cdot LW$$

where:

LA= Leaf Area as defined above in cm2

l_c = coefficient, as determined by (Clifton-Brown and Lewandowski, 2000). For *Miscanthus* it was quantified as 0.74.

LL= length of the leaf in cm measured as described above

LW= width of the leaf in cm measured as described above

The values were used to determine the rate of expansion of the leaf and the rate of the plant elongation.

6.2.2.2 Physiological measurement

6.2.2.2.1 *Photosynthetic activity*

During plant growth, the photosynthesis activity was monitored by measuring the chlorophyll fluorescence once per week on the youngest leaf with fully extended ligule as described by Ings et al. (2013). Chlorophyll fluorescence was measured between 10:30 and 12:00 with a HandyPEA continuous excitation chlorophyll fluorimeter (Hansatech Instruments Ltd., Norfolk, UK). When using the PEA, the attached leaf was dark-adapted with a leaf clip for 30 min before the measurement, sufficient to produce the opening of all the photosystem 2 (PSII) reaction centres. During the measurement, the PEA sensor unit was held over the clip, and the shutter opened (Kalaji et al., 2014b). The total saturation of all the reaction centres in PSII was then achieved with a high intensity LED array on the sensor head, providing a maximum light intensity of 3000 μmol m−2 s−1. The energy dissipation by fluorescence was recorded by the PEA, and photosynthetic parameters were automatically determined. A complete list of the parameters can be found in .

6.2.2.2.2 *Chlorophyll content*

Chlorophyll content was estimated as described by (Ings et al., 2013). It was measured on the youngest leaf with a fully expanded ligule between 10:00 and 12:00 h using a

SPAD-502 m (Konica Minolta Optics Inc.). Three readings were taken at quarterly intervals on the leaf and the mean of the values recorded.

6.2.2.2.3 *Imaging*

Images of the plants were collected 21 days after the start of the treatment (9th of April 2018). Pictures were taken using a Canon DSLR camera (EOS1100D). The camera setting was 1/80 sec. f/5 18 mm ISO400 in natural light. For all the pictures, a black non reflective background was used.

6.2.3 Sampling

6.2.3.1 Sample collection

Sixty days after the beginning of the treatment, samples for each plant were harvested. One tiller for one replicate of each genotype and each treatment was collected for future molecular biology needs. The tiller was randomly selected and its fresh weight, and that of its leaf and stem organs, was recorded. Plant samples were placed in a plastic bag and frozen in liquid nitrogen immediately after weighing. The day after, the rest of the plants was harvested. For this, all the tillers of each plant were cut at the level of the soil. Total fresh weight of the plant and the number of tillers was measured and recorded. Subsequently, each tiller was separated in leaf and stem, cut in small pieces and placed in pre-weighed, pre-labelled plastic bags. Total fresh weight of each portion, including the plastic bag, was measured and recorded. Plant material was divided into smaller pieces to fit plastic sampling bags and frozen in a freezer at -20 °C. Samples were then freeze-dried for three days until constant weight. Weight of each dried sample was measured and recorded.

6.2.4 Cell wall preparation

The cell wall was prepared, as described in subsection 3.2.1.3.

6.2.5 Matrix monosaccharide content

Monosaccharide content of the cell wall was determined as described in subsection 5.2.7.

6.2.6 Crystalline cellulose quantification

Crystalline cellulose was quantified according to (Foster et al., 2010) with some modifications according to (da Costa et al., 2018). The modification involved the use of a second acid digestion step with sulphuric acid under high temperature (120°C in the autoclave for 45 minutes) to avoid undegraded cellulose. The modifications implemented are supported by (Petit et al., 2019).

6.2.7 Saccharification efficiency

Saccharification efficiency was determined as described in subsection 5.3.6.

6.2.8 Data analysis

Data were analysed using the R programming language (R Core Team, 2018). Statistical significance between treatments was tested using an ANOVA test and a Tukey HSD test as a posthoc test.

6.3 Results

6.3.1 Phenology of the plant

6.3.1.1 Visual observation of nutrient lack symptoms

During the growth of the plant symptoms usually associated with nitrogen and phosphorus deficiency was observed in the individual treated plants. On the plants under nitrogen stress, a decrease in the green of the leaf and a size reduction in comparison with the control plants was observed. On the phosphorus limitation treatment, an increase of pigmentation in leaves and stems was observed along with an augmented size in comparison with the control condition.

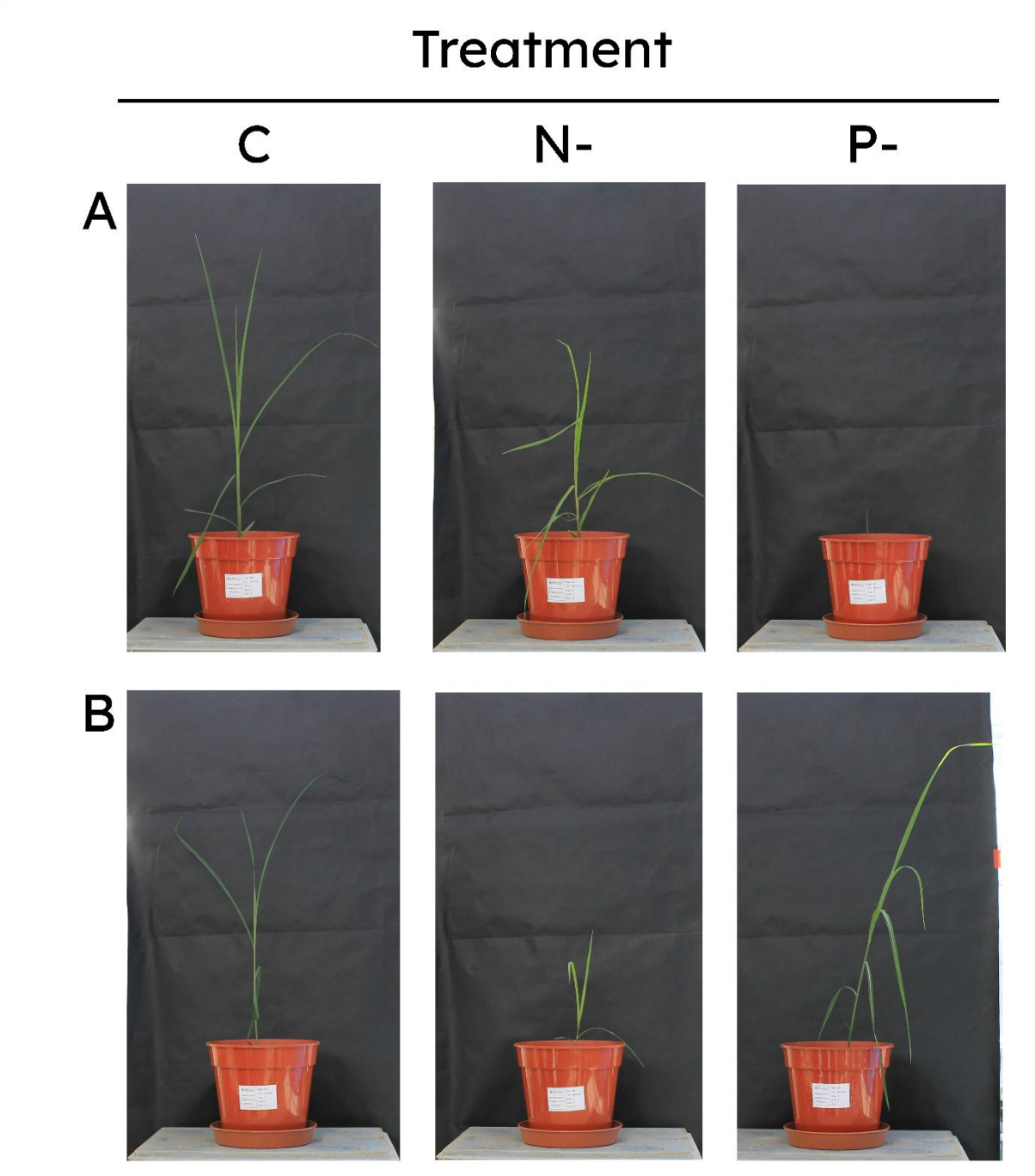


Figure 6-3 Symptoms of nitrogen and phosphorus limitation in *Miscanthus* spp. Comparison of plants of *Miscanthus* under control (C) and under nitrogen (N-) and phosphorus (P-) limitations. The phenological effects were stress and species-specific. In this picture Kay a genotype of *M. lutarioriparius* (A) and Tristan, a hybrid created at IBERS (B).

6.3.1.2 Leaf Area and expansion rate

The leaf area (LA) was measured at 10, 13, 17, 20, 27, 31, 34, 38, 41, 48, 55, and 60 days of treatment (DOT) on the highest fully expanded ligule (leaf 0). Results of the measurements are represented in Figure 6-5. For all the genotypes, differences

between treatments were starting to appear after 27 DOTs. As expected, nitrogen limitation had the most marked effect, causing a reduction of the leaf area in all the genotypes. More complex was the effect of phosphorus limitation on the leaf area. A reduction of leaf area was observed in the artificial hybrid Galahad but not in its sibling Tristan, in Kay (*M. lutarioriparius*) and Gaheris (*M. sacchariflorus*) compared with the control condition.

Conversely, phosphorus limitation increased the leaf area in Gareth (*M. sinensis* form N Japan). However, the same was not observed in the other genotypes of *M. sinensis*. Under phosphorus limitation, the 4 genotypes of *M. floridulus*, the natural hybrid and two genotypes of *M. sinensis* did not show differences in the area of the leaf compared with the control.

Overall, these results show that the effect of nutrient limitation on the leaf area depended on the specific stress and the specific genotype considered. Also, they showed that a limitation does not necessarily correspond to a reduction of the leaf area trait, maybe as a result of the activation of coping mechanisms. Since the treatments affected leaf area, an effect on the rate of leaf expansion was also expected. Surprisingly, both treatments had no effect on this parameter in any of the genotypes under study (Figure 6-5). However, it was possible to observe a tendency to an increase in the expansion rate in the initial measurements. This result may indicate that the reaction to the stresses in terms of leaf area was quick and happened in the first ten days of treatment and that after the adjustment, the rate was maintained to a level similar to the one in the control condition. Overall, it can be hypothesised that the effect of the stress on leaf area observed at the end of the experiment was the result of this initial readjustment, rather than of a progressive reduction of the leaf area.

6.3.1.3 Plant height and elongation rate

The plant heights were measured on the same dates as to when the measurements of leaf area were taken. Results of the measurements are represented in Figure 6-6. The effect of the nutrient stresses on this trait was mirroring completely the effect observed on LA. The same similarity can be observed between the rate of plant elongation and the rate of leaf expansion.

In conclusion, the results of the measurement of plant phenology suggest that nutrient stress conditions affected the overall size of the plants without affecting their proportions. Under stress conditions, biomass yield was affected mostly as a consequence of an effect on tillering, the weight of the single tiller was not affected by the stress and plants tended to maintain their architecture.

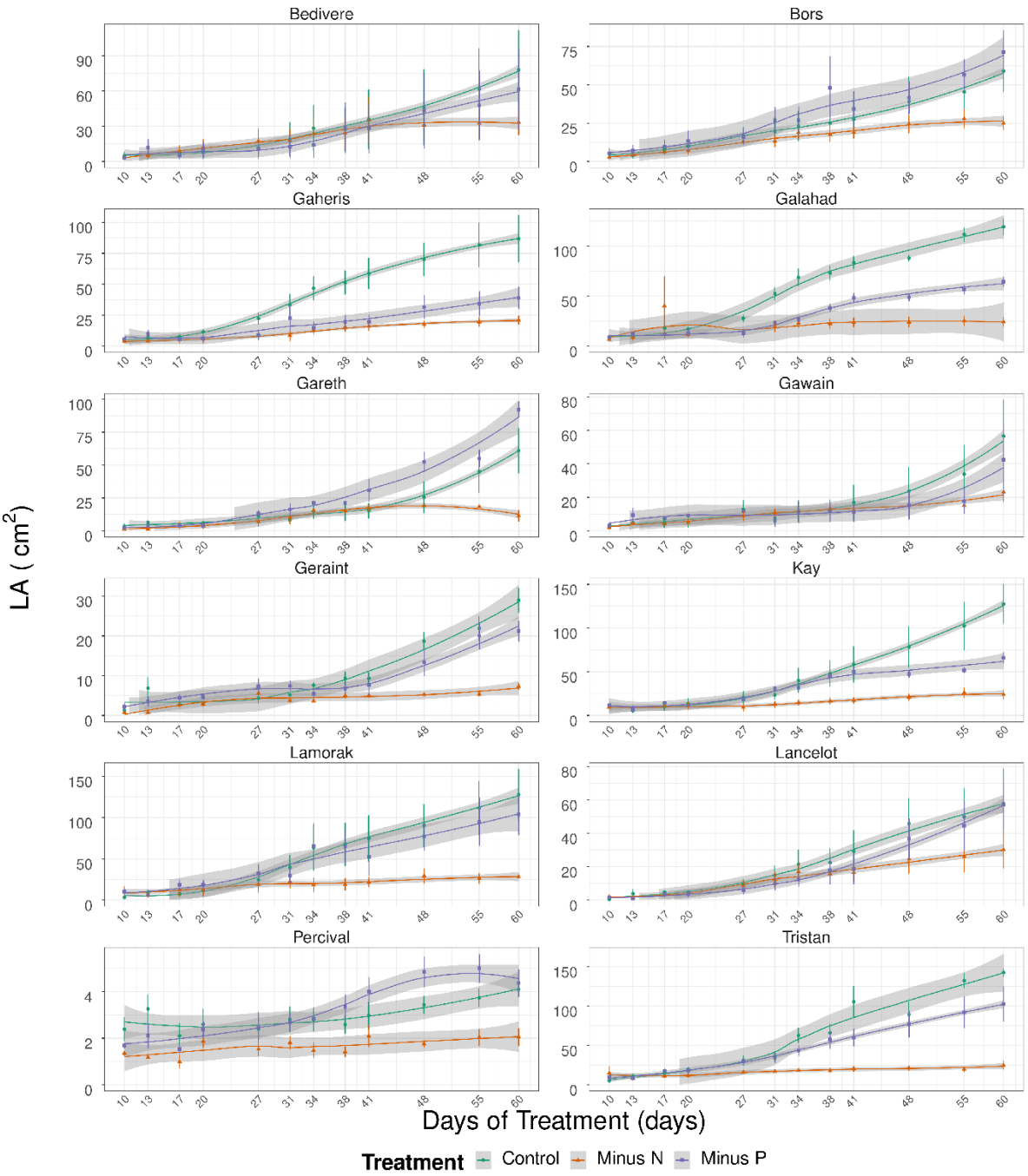


Figure 6-4 Leaf area value in 12 genotypes of *Miscanthus* under three different nutrient scenarios

The leaf area was monitored during the experiment in the three conditions imposed: control condition (Control), minimum nitrogen (Minus N) and minimum phosphorous (Minus P). Colours represent different treatments. Datapoint values are the average of measurements on 5 biological replicates of the treatment. Error bars are the standard deviation of the mean. The smoothing is performed using a Locally Estimated Scatterplot Smoothing (LOESS). The grey shaded area represents standard error of the smoothing.

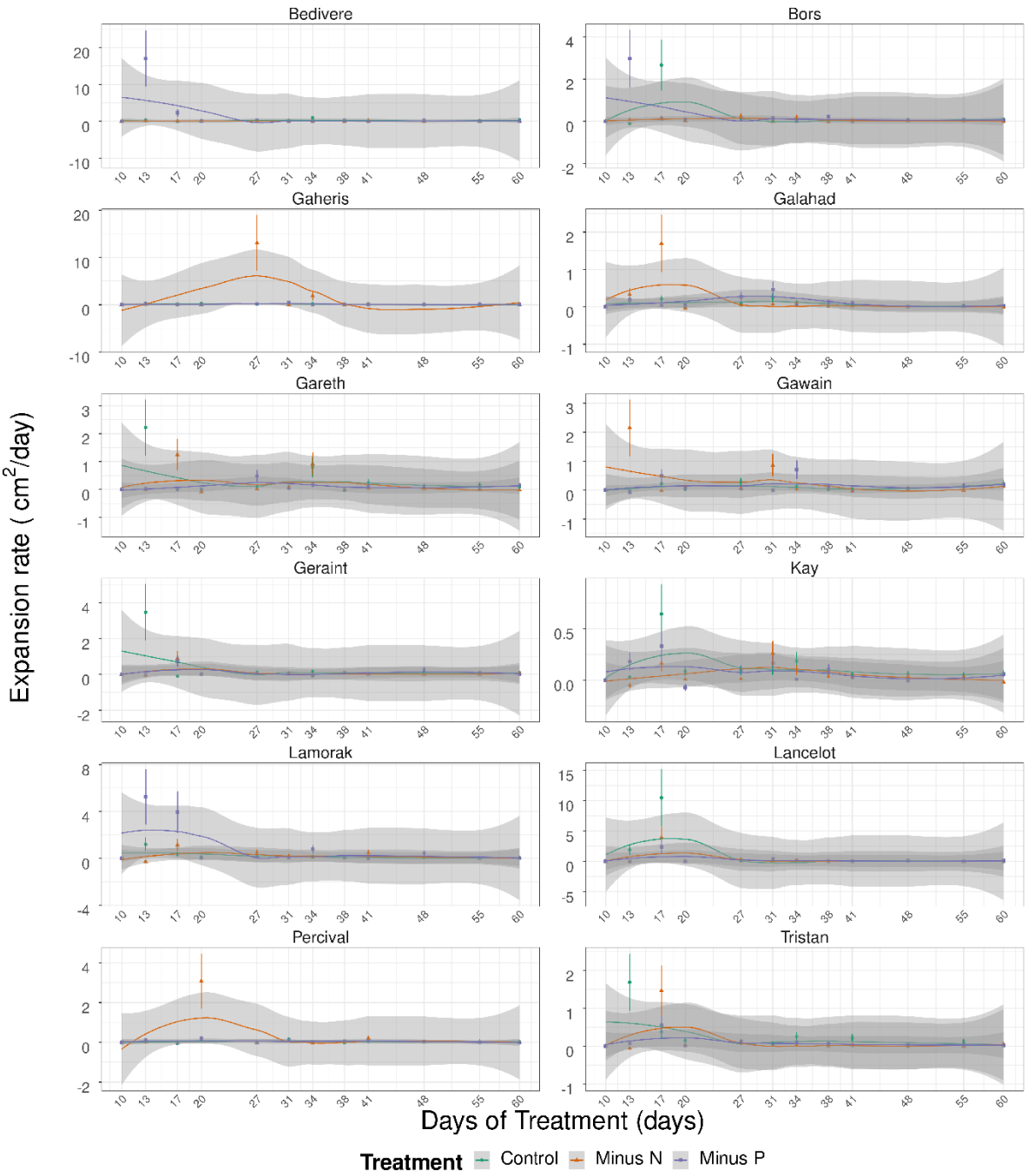


Figure 6-5 Leaf expansion rate.
Datapoint values are the average of measurements on 5 biological replicates of the treatment.

Colours represent different treatments. Error bars are the standard deviation of the mean. The smoothing is performed using a Locally Estimated Scatterplot Smoothing (LOESS). The grey shaded area represents standard error of the smoothing.

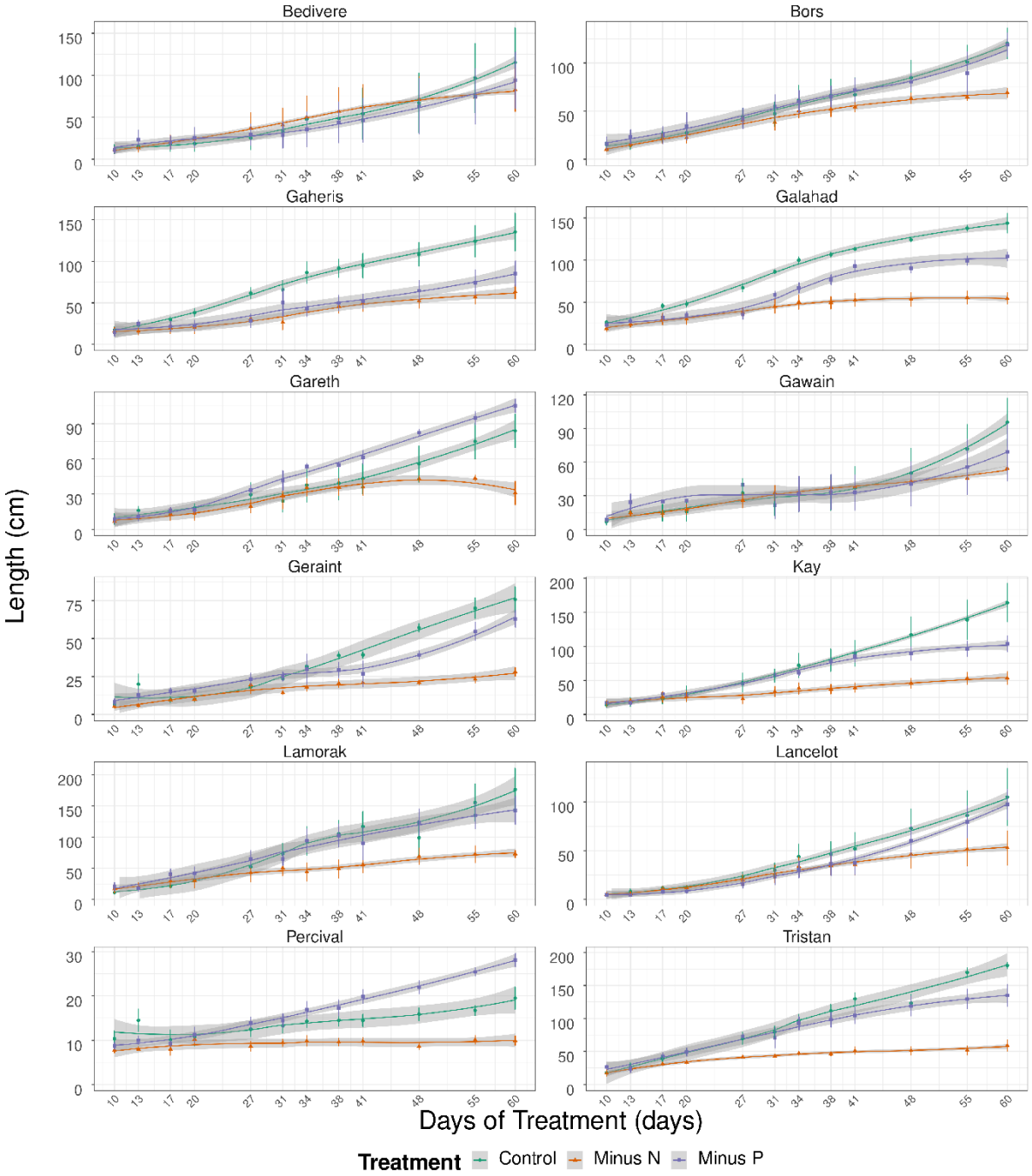


Figure 6-6 Plant height in 12 genotypes of *Miscanthus* under three different nutrient scenarios

Plant height in 12 genotypes of *Miscanthus* under three different nutrient scenarios: control condition, minimum nitrogen and minimum phosphorous. Length of the plant was measured from the pot level to the top of the youngest completely extended leaf. Datapoint values are the average of measurements on 5 biological replicates of the treatment. Colours represent different treatments. Error bars are the standard deviation of the mean. The smoothing is performed using a Locally Estimated Scatterplot Smoothing (LOESS). The grey shaded area represents standard error of the smoothing.

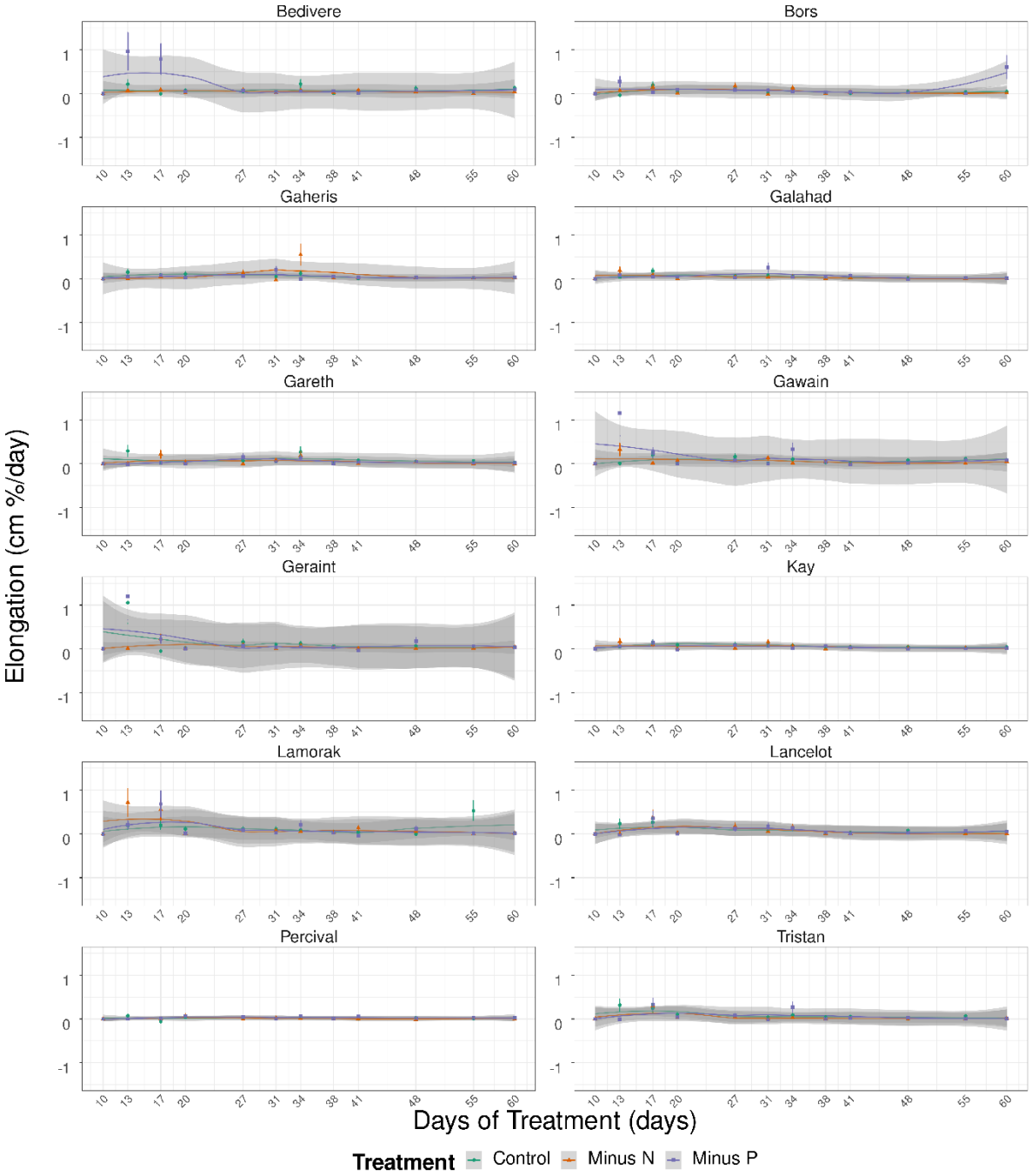


Figure 6-7 Plant elongation rate. Datapoint values are the average of measurements on 5 biological replicates of the treatment.
Colours represent different treatments. Error bars are the standard deviation of the mean. The smoothing is performed using a Locally Estimated Scatterplot Smoothing (LOESS). The grey shaded area represents standard error of the smoothing.Physiological traits

6.3.2 Chlorophyll fluorescence

6.3.2.1 JIP test

The JIP curve has been reported as an essential diagnostic tool for several nutrient stresses, including nitrogen and phosphorus limitations (Kalaji et al., 2014a; Živčák et al., 2014b; Živčák et al., 2014a; Zivcak et al., 2015; Kalaji et al., 2018b; Kalaji et al., 2018a). The JIP curves obtained from measurements at the different time point of the 12 genotypes in the three nutritional conditions were visually analysed (Figure G. 2Figure G. 3Figure G. 4). The JIP curve of the control condition was consistent between genotypes and did not change over time. The curves of genotypes under nutrient limiting conditions, both nitrogen and phosphorus, tended to have a different shape compared to the control condition. The differences were affecting all the 4 areas of the curve (O, J, I and P). However, the effect of each stress on the shape of the OJIP curve was not univocal, and it instead depended on the genotype-treatment combination considered. Phosphorus stress has been reported to affect the OJIP curve, causing a flattening of step I in maize (Carstensen et al., 2018a). Here, there was a tendency of the curves from genotypes under phosphorus limitation to maintain the shape in the I area of the curve.

6.3.2.2 JIP parameters

The analysis of photosynthetic activity resulted in a series of parameters describing the photosynthesis. The values for the parameters were automatically calculated by the Handy PEA. The Parameters are listed inFigure S 1. To simplify the set of variates, a PCA was run. Considering that a phenologic effect was evident only in the latest measurement, only the chlorophyll fluorescence data collected at 60 DOT were used for the PCA.

The PCA showed that the first two components were able to explain 75.6% of the total variability (Figure G. 5). The correlation of the variates to the first two components was measured using the square cosine (Figure G. 6) to determine the quality of representation. The variates with the highest correlation were Area, Vj, ΔV/Δt0, ΨE0, ET0/CS0, RE0/CS0, ABS/Csm, ET0/Csm, RE0/Csm. The same variates were the ones with the highest contribution to the first two components (Figure G. 7and Figure G. 8, respectively). The contribution of the variates to the components was calculated as:

Equation 6.1

$$cont^{i} . = \frac{(C1 \cdot Eig1) + (C2 \cdot Ei^{g}2)}{Ei^{g}1 + Ei^{g}2}$$

Where C1 and C2 are the contributions of the variable on PC1 and PC2, respectively, Eig1 and Eig2 are the eigenvalues of PC1 and PC2, respectively. The information about the relations between the variables and the first two components is represented by the plot of the loadings (Figure 6-8).

The plot of the PCA scores (Figure 6-8 A) shows that the first principal component was able to partially separate between photosynthetic parameters of plants grown under nitrogen stress and plants grown under control conditions. Measures of the parameters from plants grown under phosphorus limitation tended to group together with the control condition. This result seems to support the hypothesis that the limitation of phosphorus was not impacting on any of the parameters determined for the chlorophyll.

Subsequently, genotypes specific differences in the normalised area above the fluorescence transient (Sm), the slope at the origin of the fluorescence (ΔV/Δt0), the relative variable fluorescence at step J (Vj) and step I (Vi), extreme ratios (Fv/Fm) and

the performance index (PIABS) were investigated. The Sm represents the maximal area under the fluorescence curve normalised by F. It is calculated as:

Equation 6.2

$$S_m = \frac{Area}{F_m - F_0}$$

Where:

Area is calculated according to Equation 6.3

Fm is the maximum F

Fo is the F at time o

Equation 6.3

$$Area = \int_0^{t_{F_{ma}}'} (F_m - F_t)\Delta t$$

Where:

Fm is the maximum F

Ft is the F at the time t

Δt is the time difference

In all the genotypes, the Sm was not greatly affected by phosphorus limitation (Figure A 9). Under nitrogen limitation, Gaheris, Galahad, Geraint and Lamorak showed a progressive decrease in Sm during the duration of the treatment. Bors, after an initial decrease (time 3), recovered.

The ΔV/Δto showed a tendency to increase during the duration of the treatment in all the genotypes under nitrogen limitation compared with the control condition. Conversely, it was not differing significantly between genotypes under phosphorus limitation and their controls (Figure G. 10). On the following measurements Bors, Tristan, Gaheris and Percival showed a tendency to have ΔV/Δto closer to the control, which could suggest a tendency to recover.

The Vj was calculated according to Equation 6.4 (Havaux et al., 1991).

Equation 6.4

$$V_j = \frac{F_j - F_0}{F_m - F_0}$$

In the limited nitrogen treatment, all the genotypes had a Vj higher than their control conditions. Under phosphorus limiting conditions, only Galahad showed an increased Vj, while all the other genotypes did not differ significantly from the control condition (Figure G. 11).

The relative variable fluorescence at step I is defined similarly to Vj as:

Equation 6.5

$$V_i = \frac{F_i - F_0}{F_m - F_0}$$

Under phosphorus limitation, it was not significantly different from the control condition in any of the genotypes under study. At the same time, it was higher than the control under nitrogen limitation in Gaheris, Galahad, Geraint, Lamorak, Percival and Lancelot.

The Fv/Fm is the normalised ratio between the extremes (Strasser et al., 2000). Considering:

Equation 6.6

$$F_v = F_m - F_0$$

Then:

Equation 6.7

$$\frac{F_v}{F_m} = \frac{F_m - F_0}{F_m}$$

Both the nutrient-limited conditions tended a ratio in the range between 0.65 and 0.75, lower than the control (~0.8) at the intermediate (3 to 5) times of the experiment.

The ratio was tendentially lower under nitrogen limitation compared with the phosphorus limitation. There was a tendency of a recover in most of the genotypes at the latest times. The ratio in Galahad, Geraint and Kay tended to drop in the control condition at the last measurement. This could be attributed to a slight water deficit. It cannot be excluded that the bigger size of the plants of this genotypes, resulted in higher water consumption of the plants in the last week of the experiment (Figure G. 12).

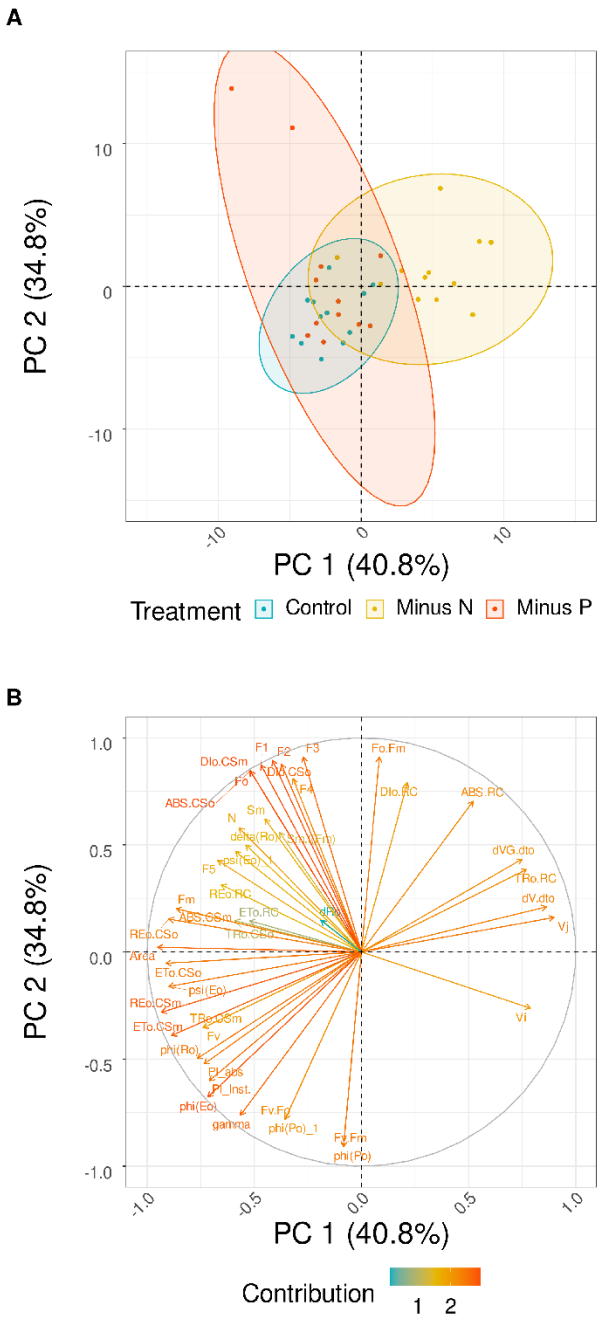
The PIABS is comprehensive of several parameters of the photosynthesis above. It is defined by Equation 6.8 (Živčák et al., 2014b).

Equation 6.8

$$\boldsymbol{PI_{AB}}^{\boldsymbol{S}} = \frac{1 - \frac{F_0}{F_m}}{\frac{M_0}{V_j}} \times \frac{F_m - F_0}{F_0} \times \frac{1 - V_j}{V_j}$$

The PIABS was consistently lower in all the genotypes under nitrogen limitation and tended to decrease during the experiment compared with the control condition. Under phosphorus limitation, only Galahad and Lamorak tended to have lower PIABS compared with their controls, and the difference increased in time, while they were not different compared with the nitrogen limitation condition.

Taken together, the results on the effect of nutrient stress on the parameters of chlorophyll fluorescence indicate a strong effect of nitrogen on different *Miscanthus* genotypes grown under nitrogen limitation, a limited effect of phosphorus limitation on the same parameters and a different tendency to recover from the stress between genotypes.



Parameter	Description
t_for_Fm	Time to Fm
Area	Area under curve
Fo	Minimum value for chlorophyll fluorescence
Fm	Maximum value for chlorophyll fluorescence
Fv	Variable fluorescence
F1	Value of fluorescence at O
F2	Value of fluorescence at J
F3	Value of fluorescence at I
F4	Value of fluorescence at P
F5	
Fo.Fm	Ratio between the extremes expresses the amplitude of the ChlF transient rise
Fv.Fm	Maximum quantum efficiency of PSII photochemistry.
PI_Inst.	Instantaeous Photosynthetic Index
Fv.Fo	Ratio Fv/F0
Vj	Relative variable fluorescence at the J-step
Vi	Relative variable fluorescence at the I-step
dVG.dto	dVG.dto
dV.dto	dV.dto
Sm	Normalized area above the transient.
N	Number of turnovers of QA
Sm.t(Fm)	Normalized area above the transient at Fm
ABS.RC	Absorption flux (of antenna Chls) per RC
Dlo.RC	Dissipated energy flux per RC (at t = 0)
TRo.RC	Trapping flux (leading to QA reduction) per RC
ETo.RC	Electron transport flux (further than) per RC
REo.RC	Electron flux reducing end electron acceptors at the PSI acceptor side per RC
phi(Po)	Maximum quantum yield of primary photochemistry (at t = 0)
psi(Eo)	Probability an exciton moves an electron beyond Qa
phi(Eo)	Quantum yield of electron transport (at t = 0)
delta(Ro)	Prob.that an electron from the intersystem reduces electron acceptors at the PSI end
phi(Ro)	Quantum yield for reduction of end electron acceptors at the PSI acceptor side
ABS.CSo	Absorption flux for cross section at 0
Dlo.CSo	Dissipated energy flux per cross section at 0
TRo.CSo	Trapped energyflux (leading to QA reduction) per cross section at 0
ETo.CSo	Electron transportflux per cross sectionat 0
REo.CSo	Electronflux reducing end electron acceptors at the PSI, per cross section at 0
ABS.CSm	Absorption flux for cross section at Fm
Dlo.CSm	Dissipated energy flux per cross section at at Fm
TRo.CSm	Trapped energyflux (leading to QA reduction) per cross section at at Fm
ETo.CSm	Electron transportflux per cross sectionat at Fm
REo.CSm	Electronflux reducing end electron acceptors at the PSI cross section at at Fm
gamma	gamma
phi(Po)_1	aximum quantum yield for primaryphotochemistry
psi(Eo)_1	Probability (at t = 0) that a trapped exciton moves an electron QA
PI_abs	Performance index
dRo	Probability that an electron is transferred to reduce endelectron acceptors at the PSI
PI_total	Performance index
DF_abs	Delayed (chlorophyll) fluorescence Absolute
DF_Total	Delayed (chlorophyll) fluorescence total
kPABSkF	kPABSkF
kNABSkF	kNABSkF
Origin_to_F1	Origin to F1
F1_to_F3	F1 to F3
F1_to_F4	F1 to F4
F1_to_F5	F1 to F5
F3_to_F4	F3 to F4
F4_to_F5	F4 to F5
F5_to_Fm	F5 to Fm

Figure 6-8 Loadings of the variates on the first two components.
The length of the arrows is proportional to the contribution of each variate to the components. The angle between the arrows and the axes describes the quality of representation of the component for each variate. The colour is proportional to the contribution of each variable to the components.

6.3.3 Chlorophyll content

The nutrient stress could have impacted on the chlorophyll content. It has been shown that limited nutrient availability is associated with a reduction of photosynthetic pigment (Okunlola et al., 2014;). In addition, the effect of nutrient limitation can be differential between varieties of the same species. For example, levels of chlorophyll content in varieties of maize tolerant to low nitrogen were less affected by nitrogen limitation compared with non-tolerant varieties of the same species (Wu et al., 2019). There is no information available for *Miscanthus* on this regard. However, considering that the species used in this experiment came from soils that potentially had different nutrient availabilities, it was hypothesised that the genotypes had different tolerances to nutrient limitation. Consequently, an effect of the genotype and of the treatment on the content of chlorophyll was expected.

The amount of chlorophyll was assessed using a SPAD meter (Ansatech) on the youngest completely expanded leaf at 4 time-points during plant growth. The fourth time point being the measurement the day before the harvesting and the first measurement 10 DOTs. Without a calibration for each species of *Miscanthus* in this experiment, the SPAD did not provide absolute quantification of the chlorophyll contained in the organ tested, the values are here used only for relative comparison between plants under treatment and control conditions. Results of the measurements are shown in Figure 6-9. The large error bars are explained by a large variation observed between biological replicates. The amount of chlorophyll tended to increase in the control condition of all the genotypes during the experiment. In Gaheris, Galahad, Gawain, Geraint, Lamorak, Lancelot ant Tristan, a possible tendency to plateau close to harvesting can be observed.

All the genotypes under nitrogen limiting conditions had a lower amount of chlorophyll, and the difference increased along with the number of DOTs. In all the genotypes, the amount of chlorophyll under nitrogen limiting conditions was significantly different ($p< 0.05$) from the one under control conditions (Figure 6-6).

in the latest measurement the amount of chlorophyll by SPAD , was mostly lower under phosphorus limiting conditions, than under control conditions in Bedivere (27%), Bors (31%), Gaheris (19%), Galahad (42%), Kay (53%), Lancelot (47.2%), Percival (54%) and Tristan (29.02%).. In Gawain and Lamorak, at the last measurement, the amount of chlorophyll under phosphorus limitation was not significantly different from the one in its control. In Lamorak the amount of chlorophyll in phosphorus limiting conditions compared to the control condition increased at a higher rate until the second time point, on the third time point there was no significant difference between control and phosphorus limitation.

Overall, the measurements of chlorophyll content showed that the nutrient stress affected the amount of chlorophyll in all the genotypes in the earliest time points. While the lack of chlorophyll was not reversed under nitrogen limiting conditions, under phosphorus limitation, there was a higher rate of increase of chlorophyll content that allowed some genotypes to recover to content not different from the control condition. Chlorophyll is essential to the photosynthetic process. Its reduction under nitrogen limitation could be a consequence of the central role of nitrogen in the structure of chlorophyll, which may cause a reduction of chlorophyll synthesis for lack of substrate from which the plant cannot recover. Conversely, under phosphorus limitation, the initial lack of chlorophyll could have resulted from a limitation in energy in the chlorophyll synthetic pathway due to lack of phosphorus for the ATP, that was eventually overcome by change the priority of other pathways.

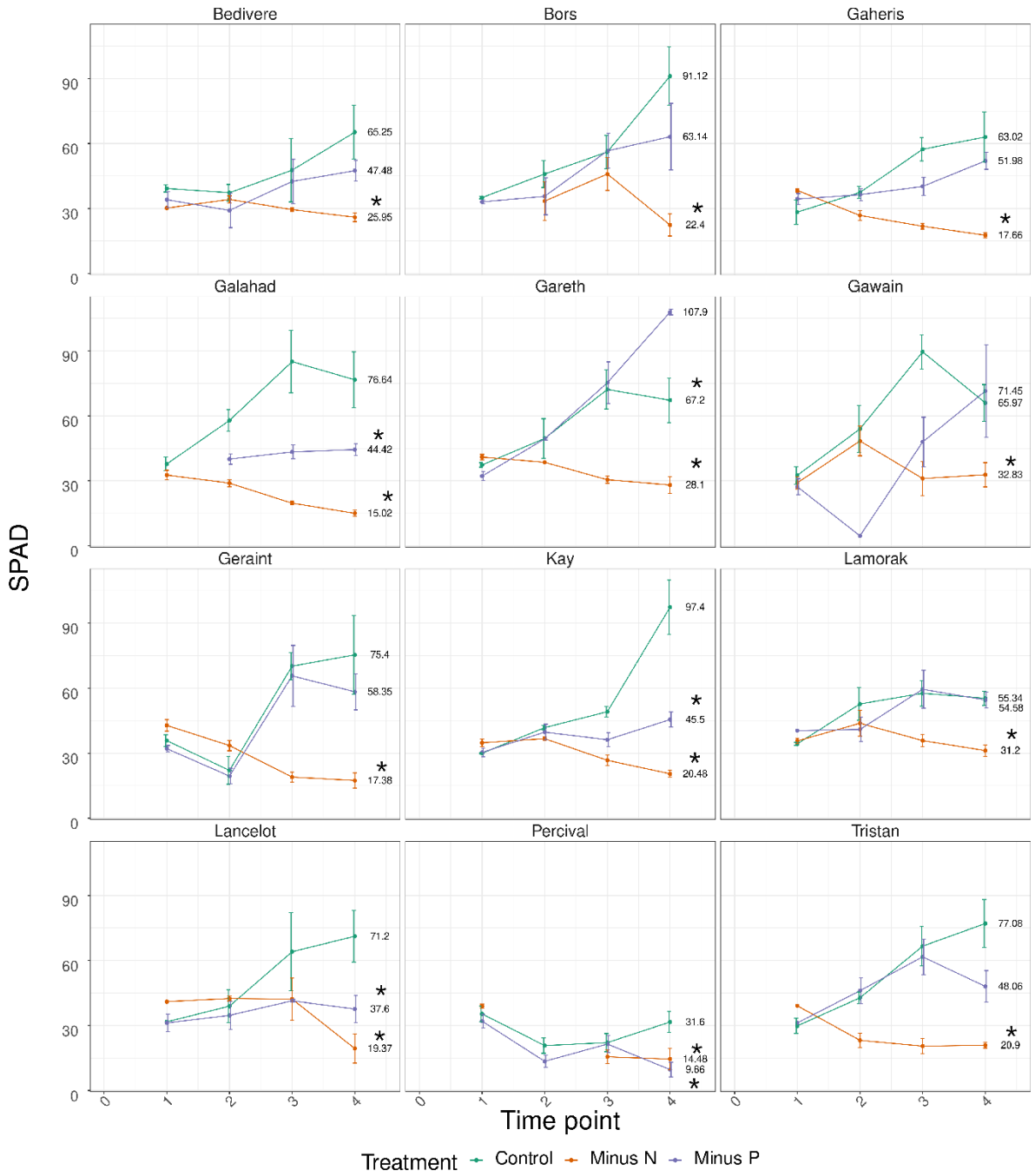


Figure 6-9 I Chlorophyll content on the youngest completely expanded leaf of the 12 genotypes.
The amount of chlorophyll was quantified using a SPAD meter (Hansatech). Values on the y-axes are the chlorophyll content as the average of three measurements in the points of each leaf and on at least 3 biological replicates. Bars represent the standard deviation of the mean. Values on the x-axes are times of the measurement. The first measurement was carried 10 DOTs. Colours represent the treatment.

6.3.4 Biomass yield

6.3.4.1 Plant survival

After 60 days of treatment alive and dead plants were counted. Dead plants were those that after transplant stopped growing. The pot in the case of a dead plant contained just soil. Accordingly, the percentage of survival calculated. Results of the average by genotype, genetic group and treatment are shown in

Table 6-7 . The plant survival varied between the genetic groups in the experiment between 46.67±11.55% and 93.33±10.33 %. Only *M. sinensis* from North Japan and artificial hybrid (Hyb.gig) groups were significantly different from each other.

Table 6-7 Percentage of plants alive at the end of the experiment.

Genetic group	Survival (mean±sd)
	66.67±21.46
Flor	ab
Hyb.gig	86.67±11.55 ab
Hyb.mx	93.33±10.33a
	53.33±23.09
Lut	ab
Sacc_rob	93.33±11.55 ab
Sin_N_Japan	46.67±11.55 b
Sin_S_Japan	73.33±11.55 ab
	86.67±23.09
Sin_Taiwan	ab

6.3.4.2 Biomass production

At the end of the experiment, the total biomass of the plants was harvested and freeze-dried. The number of tillers and the total dry weight of each plant, along with the dry weight of leaf and stem material, was recorded. The average weight of the single tillers was calculated as a ratio between the total dry weight and the number of tillers. Differences in these variates between genotypes, genetic groups and treatment were investigated with an ANOVA test (Fisher, 1919). Summary values (mean±sd) for dry weight production (DW), number of stems (n_stems) and percentage of leaf material are reported Table G. 1 Dry weight of plant, stem and leaf and relative leaf content. Results of the test of significance are reported in Table G. 8.

When the total amount of dry biomass yield is considered (Table G. 1), an effect of both stresses can be observed for two genotypes (Galahad, Tristan) and high dispersion of the values around the median. However, the normalisation of the dry weight value for the number of tillers reduced this dispersion and showed that there was no significant effect of the treatment on the weight of single tillers (Figure 6-2), while the effect was significant for the number of tillers (Table G. 8. Nitrogen limitation resulted in a statistically significant reduction of tillering in all the genotypes compared with the control condition. Under phosphorus limitation, only Galahad (a hybrid) and Gaheris had a significantly lower number of tillers. Comparing the nitrogen limitation and phosphorus limitation conditions, the number of tillers was different in the two conditions only for the *M. floridulus* genotype, Geraint.

Taken as a whole, these results show that nitrogen stress in the 12 genotypes of *Miscanthus* reduced biomass production by affecting the tillering. This is could be related to the typical reaction of perennial plants to stressful conditions. It is known that, differently from annual plants, perennial plants, in nutrient limiting conditions

are able to efficient prevention of nutrient losses by reducing their growth (Rennenberg et al., 2010)

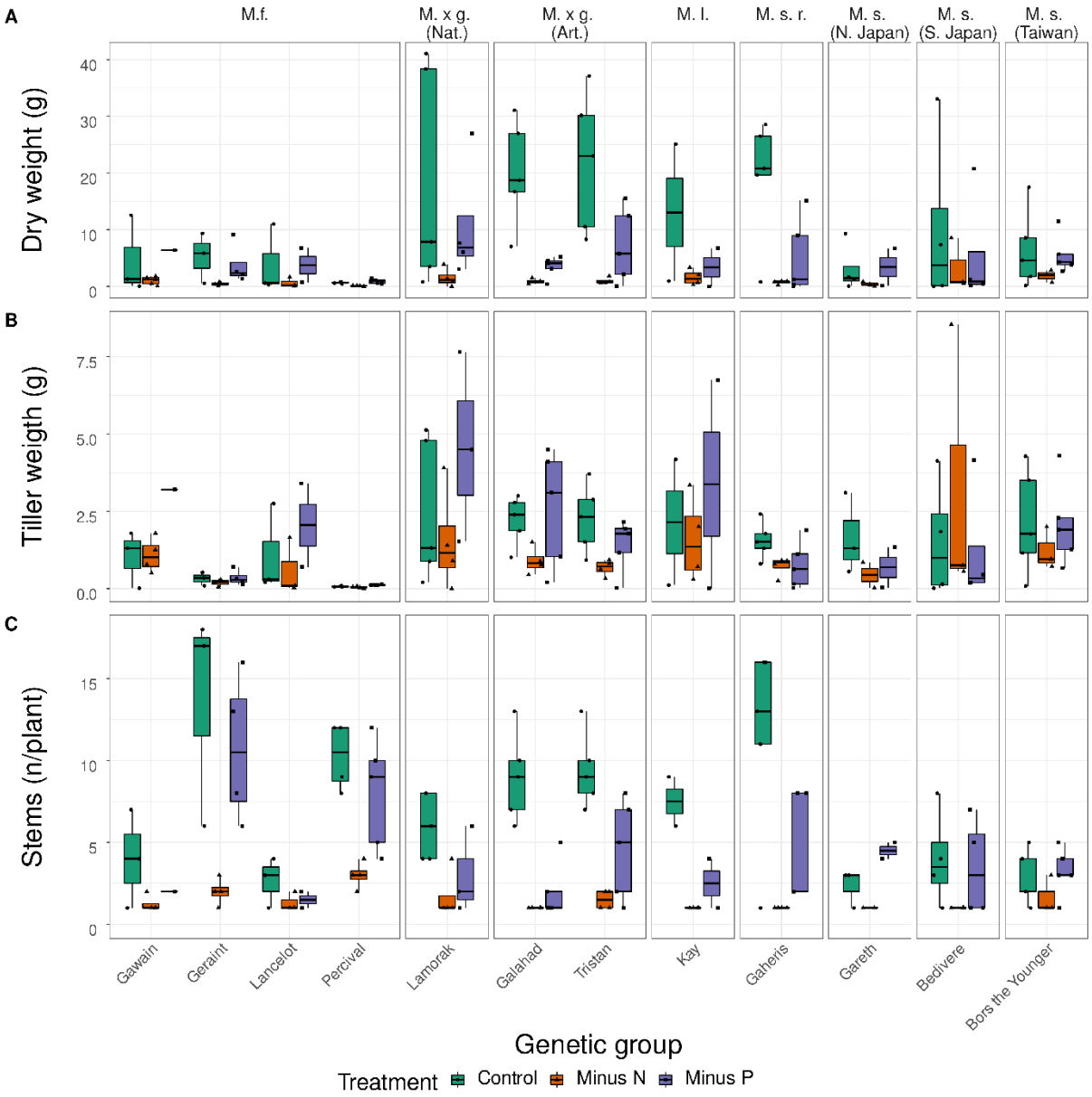


Figure 6-10 Effect of genotype and treatment on the components of yield
The effect of genotype on biomass yield (A), tiller weight (B) and number of tillers per plant (C) was investigated. The nitrogen limitation (orange) had a significant effect only on the number of tillers in all the 12 genotypes compared with the respective control condition ones. Values are the average of the biological replicates. Under phosphorus limitation, only Galahad (a hybrid) and Gaheris had a significantly lower number of tillers compared with their controls. Boxes represent the mid two quartiles with the median drawn. The colour code represents the treatments. Control (green) receiving a full Hoagland solution. Minus N (orange) treated with a Hoagland solution without nitrogen. Minus P (purple) treated with a Hoagland solution lacking phosphorus. Different letters represent statistically significant values, according to Tukey HSD ($\alpha = 0.05$).

6.3.4.3 Leaf percentage

The different number of tillers could affect the ability to compare the proportion between leaf and total biomass between treatments and genotypes. Thus, the percentage of the leaf over the weight of the plant was normalised for the number of tillers. There was significant effect of genotype and genetic group on the percentage of the leaf,. The treatment had no significant effect and plants under nutrient limitation were maintaining the structure of each tiller..

6.3.5 Biomass quality

6.3.5.1 Saccharification efficiency

Plant biomass was harvested at a stage comparable to the principal growth stage 2, substage 9 of the BBCH scale for *Miscanthus* × giganteus morphological development stages according to Tejera et al. (2017). . However, the nutrient limitation caused a dramatic reduction in size and amount of biomass production (section 6.3.3). As a consequence, the quality for biomass conversion could be analysed only on leaf samples of 3 biological replicates in Bors, Gaheris, Galahad, Lamorak and Tristan that had enough material available.. .

The quality for the conversion of the biomass from the plants under control and nutrient limiting conditions was assessed with a saccharification test. Biomass was digested for 72 hours with a commercially used enzymatic cocktail. After the digestion, the amount of arabinose, glucose and xylose released in the supernatant was quantified and the amount, normalised by the total amount of each monosaccharide present in the cell wall, used as an indicator of conversion features. The experiment was performed in 4 technical replicates, and the values averaged. Results for each genotype and condition are shown in Table G. 3. To account for the different number of replicates available for each sample, differences were statistically tested using a

Type III Analysis of Variance test (ANOVA)(Fisher, 1919; Herr, 1986; Langsrud, 2003), and located using an HSD posthoc test (Tukey, 1949). The result of the statistical analysis is graphically presented in Figure 6-11. The nutrient stress affected the amount of all the three monosaccharides released in the supernatant.

The omnibus ANOVA test showed a significant interaction between genotype and treatment for the amount of arabinose, glucose and xylose released. This means that biomass from genotypes under treatment differed from the one of those under control condition according to the treatment genotype combination considered (Table G. 9). Therefore the effect of the treatment on the monosaccharide release was tested separately for the genotypes (Table G. 10 and Table G. 11).

The nitrogen limitation affected the amount of arabinose and glucose released in Gaheris and on the amount of all the three monosaccharides released in Galahad (Figure 6-11). However, the effect was different in the two genotypes. Gaheris showed a significant increase of arabinose released ($p = 0.036$) and a slightly significant increase of glucose released ($p = 0.071$) compared with its control condition. Galahad showed a significant decrease in the amount of arabinose ($p = 0.029$), glucose ($p = 0.029$) and xylose ($p = 0.029$) compared with the respective controls.

The phosphorus limitation resulted in an effect on the amount of all the three monosaccharides released in Gaheris and Tristan and on the amount of xylose released in Galahad. Similarly, to nitrogen stress, the effect was contrasting between genotypes. In Gaheris under phosphorus limitation, there was an increased amount of arabinose (ANOVA, $p = 0.035$), glucose (ANOVA, $p = 0.011$) and xylose (ANOVA, $p = 0.048$) released compared with the control. At the same time, there was a significantly lower release of arabinose (ANOVA, $p = 0.035$), glucose (ANOVA, $p = 0.006$) # in Tristan under phosphorus limitation compared with its control condition. In Galahad under

phosphorus limitation, the amount of arabinose and glucose enzymatically released from the biomass was not significantly different from that of the same genotype under control conditions.

In contrast, the amount of xylose decreased significantly (ANOVA, $p = 0.007$). In only one occasion, a significant difference between the two nutrient limiting conditions was observed. Tristan under phosphorus limitation had a lower amount of glucose release compared with its counterpart under nitrogen limitation.

Turning to the significant effect of genotype (Figure 6-11**Error! Reference source not found.**), it can be noticed that genotypes varied according to the nutritional condition considered. The nitrogen limitation stress induced a variation of saccharification efficiency that reduced the extent of the variance of the saccharification traits between the genotypes, while the phosphate limitation stress increased the variability of these traits.

Taken together the results of this experiment showed that nutrient limitation could affect the recalcitrance of biomass to conversion in the group of genotypes tested and that the effect is variable according to the genotype considered and the specific stress imposed. The behaviour could be attributed to different factors impacting the recalcitrance of biomass, including cell wall composition as well as linkages between components. The monosaccharide total and relative abundance in the biomass of the genotypes were quantified, and results are described in section 6.3.7 6.3.6, 6.3.7 and 6.3.8.

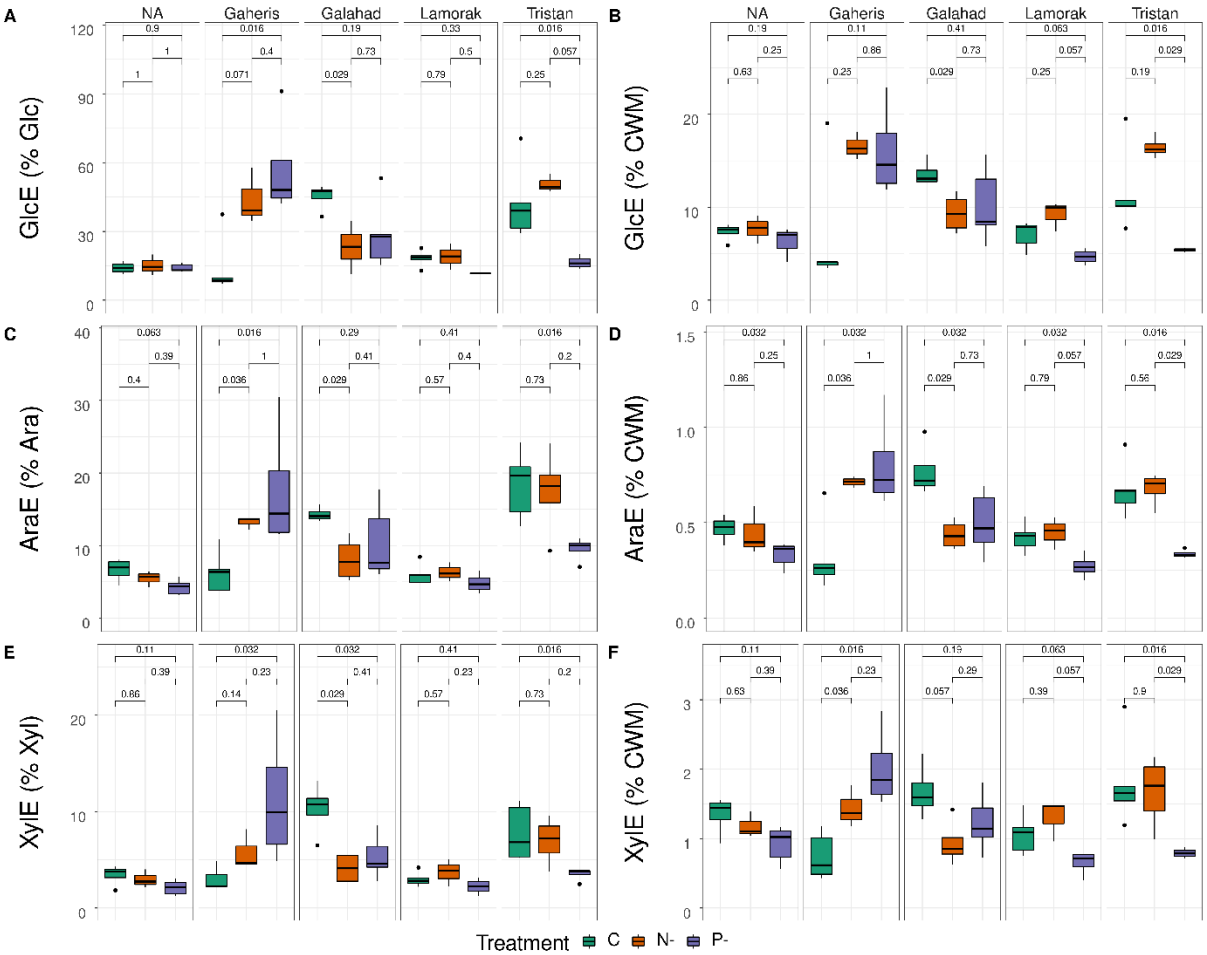


Figure 6-11 Effect of treatment on the amount of monosaccharide enzymatically released.

The treatment had an effect on the amount of monosaccharides released that was interacting with the genotype considered. Consequently, the effect of treatment was tested on the single genotypes separately using an ANOVA test of significance followed by a Tukey post-hoc test. Biomass was digested digestion with 72 hours digestion with the CellTech enzymatic cocktail. Values are the average of four technical replicates on at least 3 biological replicates of each genotype in each condition. The amount is expressed as a percentage in weight of each monosaccharide out of the total monosaccharide present in the cell wall. Boxes represent the mid two quartiles with the median drawn. The colour represents the treatment. Control (blue) receiving a full Hoagland solution. Minus N (yellow) treated with a Hoagland solution without nitrogen. Minus P (orange) treated with a Hoagland solution lacking phosphorus. Different letters represent statistically significant values as determined by the Type III ANOVA and the Tukey HSD test ($\alpha = 0.05$).

6.3.6 Total monosaccharide content

On the same biomass used for the measurement of the saccharification efficiency, the total amount of monosaccharides (MnsT) as a percentage of the cell wall weight was quantified. The MnsT was calculated by adding up the amount of each monosaccharide in the matrix and from the cellulose. MnsT represents the portion of the cell wall that is composed by carbohydrates. The effect of genotypes and nutrient stress was tested using a Type III ANOVA and a Tukey HSD test.

The preliminary two-way ANOVA revealed a significant interaction between genotype and treatment (ANOVA, p =0.00435). Therefore, the effects of these two factors were tested separately on the levels of the other factor with separate ANOVA tests.

Nutrient limitation resulted in significant differences between treatments in Gaheris (ANOVA, p = 0.011) and Galahad (ANOVA, p = 0.022). Under nitrogen and phosphorus limitation, Galahad had a significantly higher percentage of monosaccharides in the cell wall compared with its control, with the two treatments not resulting in significant differences (Figure 6-12). Under phosphorus limitation, but not under nitrogen limitation Gaheris had a significantly lower content of monosaccharides in the cell wall (t-test, p = 0.016).

Genotypes under nutrient stress conditions were ranking differently from the ones under control conditions (Figure 6-12). Bors stands out in a separate group in all the three treatments. The differences between Gaheris and Galahad as well as between Lamorak and Tristan that are significant in the control condition, under nitrogen and phosphorus limitation are reduced and resulted not significant. Under control conditions, Lamorak is significantly different from Bors in terms of total monosaccharides in the cell wall. The two genotypes have not significant differences in the trait under nitrogen and phosphorus limitation, although the increase of total

monosaccharides in Lamorak under nutrient limitation is not significant when it is compared with its control,

Overall, the effect of nutrient stress on the total amount of monosaccharide content in the cell wall is specific according to the genotype considered and the stress. The same stress can affect positively one genotype and negatively another. These differences result in the nutrient stress affecting the reciprocal differences between genotypes and impact the ranking. This could result from the genotypes having a different tolerance to nutrient limitation as a consequence of adaptation of different nutrient availabilities in their natural environment.

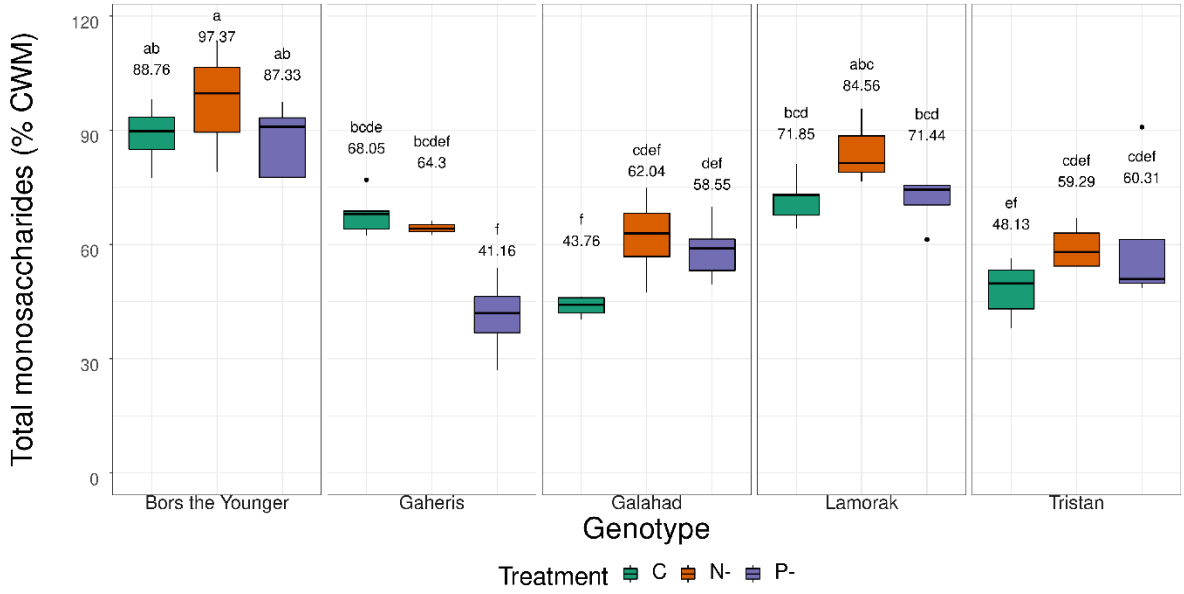


Figure 6-12 Effect of treatment on total amount of monosaccharides in the cell wall.
The total amount of monosaccharides was determined by adding up monosaccharides from hemicellulose and monosaccharides from cellulose. Values are the average of four technical replicates on at least 3 biological replicates of each genotype in each condition. The amount is expressed as percentage in weight of the cell wall. Boxes represent the mid two quartiles with the median drawn. The color represents the treatment. Control (blue) receiving a full Hoagland solution. Minus N (yellow) treated with a Hoagland solution without nitrogen. Minus P (orange) treated with a Hoagland solution lacking phosphorus. Different letters represent statistically significant values as determined by the Type III ANOVA and the Tukey HSD test ($\alpha = 0.05$). Top, plot of the differences between different treatments on the same genotype. Bottom, plot of the differences between genotype under the same nutritional condition.

6.3.7 Matrix monosaccharides content

Matrix monosaccharides are the monosaccharides contained in the hemicellulose fraction of the cell wall. They were released upon digestion with TFA. The monosaccharides determined were fucose (Fuc), arabinose (Ara), glucose (Glc), xylose (Xyl), mannose (Man), and cellobiose (Cel). A summary of the quantification is presented inTable G. 4 and Table G. 5. The differences between genotypes and treatments were tested using a Type III ANOVA test and a Tukey HSD test (α=0.05). The differences in arabinose, glucose and xylose content between treatments and genotypes are reported in Figure 6-13. There was a significant difference in the amount of glucose in the cell wall matrix while there were no significant differences detected between treatments and control in the amount of fucose, arabinose, mannose and cellobiose. Under nitrogen limitation there was a significant increase in the amount of glucose from the matrix, while the amount was not significantly different under phosphorus limitation.

The effect of treatment different in extent and direction between genotypes.. For example, nitrogen limitation tended to reduce the variability on the hemicellulose glucose content compared with the control condition (Figure 6.11 A). Considering that cell wall matrix is synthesised using nucleotide monosaccharides , what observed in this experiment could relate to the fact that nitrogen has a central role in the monosaccharide nucleotide production.

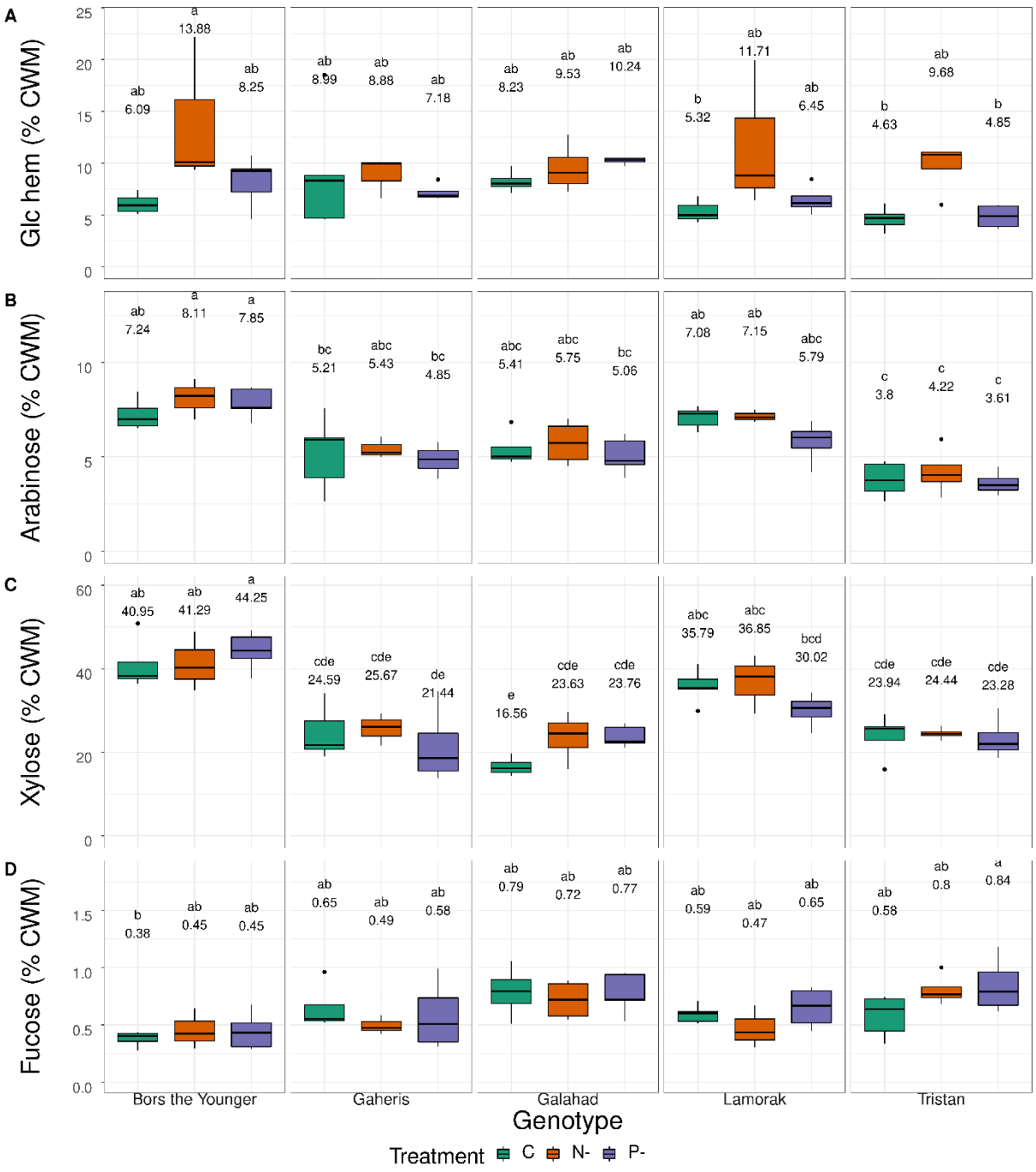


Figure 6-13 Effect of genotype and treatment on matrix monosaccharides.
Genotype and treatment (both nitrogen limitation and phosphorous limitation had a significant effect on the in the amount of glucose (Glc hem, A), arabinose (Ara, B), xylose (Xyl, C) and fucose (Fuc, D), in the cell wall matrix expressed as a percentage of the cell wall material (CWM) . Values are the average of four technical replicates on at least 3 biological replicates of each genotype in each condition. The amount is expressed as percentage in weight of the cell wall. Boxes represent the mid two quartiles with the median drawn. The color represents the treatment. Control (blue) receiving a full Hoagland solution. Minus N (yellow) treated with a Hoagland solution without nitrogen. Minus P (orange) treated with a Hoagland solution lacking phosphorus. Different letters represent statistically significant values as determined by the Type III ANOVA and the Tukey HSD test (α = 0.05). Top, plot of the differences between different treatments on the same genotype. Bottom, plot of the differences between genotype under the same nutritional condition.

6.3.8 Crystalline cellulose and hemicellulose contents

After the determination of hemicellulose composition, the cellulose was extracted from the remaining pellet using the Updegraff method (Updegraff, 1969). The cellulose was then digested twice with sulphuric acid, and the amount of glucose yielded used to quantify the amount of crystalline cellulose (CrCel). The test was performed in 4 technical replicates and values averaged. The summary of the result of the quantification is shown in Table G. 6. Differences were tested for statistical significance using a Type III ANOVA test and an HSD Tukey for posthoc analysis. Results of the statistical analysis are reported in Table G. 12, Table G. 13 and Table G. 14 and graphically presented in Figure 6-14. The effect of the interaction between genotype and treatment on the amount of crystalline cellulose in the cell wall was significant (ANOVA, $p = 0.011$). Therefore, the effect of each factor was tested separately on the levels of the other factor.

None of the genotypes tested under nitrogen limiting conditions had a significantly different amount of CrCel when compared with their controls. Under phosphorus limitation, the CrCel was significantly lower in Gaheris in comparison with its control (ANOVA, $p = 0.03$ and Tukey HSD, $p<0.05$).

Under control and phosphorus limiting conditions, Bors had a significantly higher content of crystalline cellulose compared with the other genotypes (Tukey HSD corrected). Under control conditions, the remaining genotypes cluster in two groups. Gaheris and Lamorak had a significantly higher content of crystalline cellulose in their cell walls compared to Galahad and Tristan. Under phosphorus limiting condition, there were no significant differences between the two groups. Tristan and Lamorak,

the two sibling hybrids, differed significantly for the crystalline cellulose under control conditions were not significantly different under nitrogen and phosphorus stress.

The nutrient stress had an impact on the amount of hemicellulose in the cell wall of Gaheris, Galahad and Lamorak. However, the Tukey test showed significant differences to an alpha 0.05 only in Gaheris and Lamorak. In both these genotypes, the content of hemicellulose in the cell wall under nitrogen limitation condition was not significantly different from the one under control condition. Under phosphorus limiting conditions, both genotypes had a significantly lower content of hemicellulose in their cell walls.

Overall, the nutrient stress was altering the composition of the cell wall of the genotypes resulting in alterations of the biomass quality. Such alterations could nullify the superior quality of one hybrid compared to the other or between genotypes of the same species, for example. This implies that the screening of new hybrids has to include the flexibility of their characteristics under abiotic stress. Bors seems a resilient genotype in terms of the content of cell wall components to stress. It withstood the stressful condition imposed and maintained the higher content of crystalline cellulose under both the nutrient limiting conditions imposed.

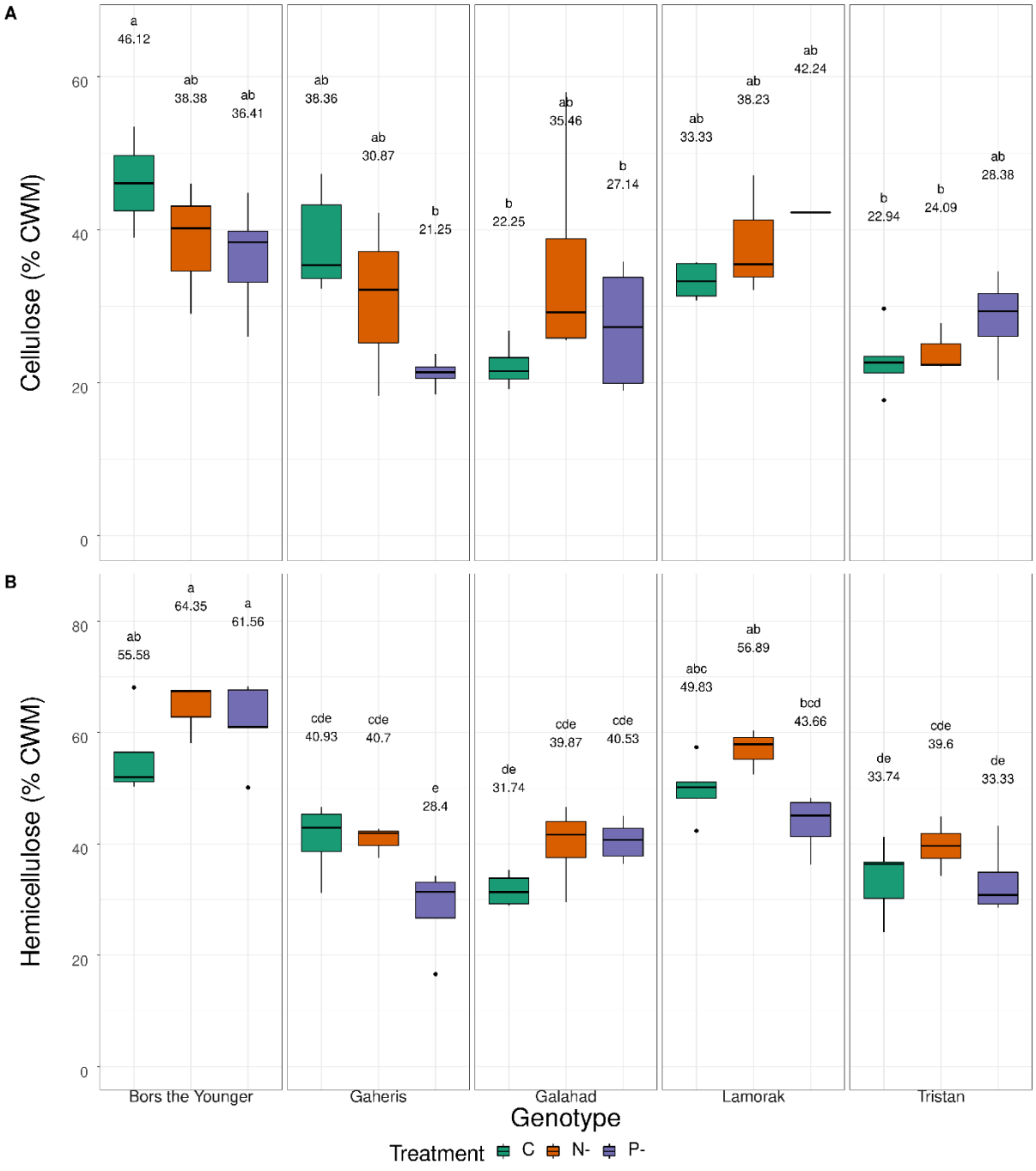


Figure 6-14 Effect of treatment on the amount of hemicellulose in the cell wall in the leaf material of the 5 genotypes.
The limitation of nitrogen and phosphorous both had a significant effect on the amount of cellulose and hemicellulose in the cell wall from leaf material of the 5 genotypes tested. The effect was interacting with the genotype considered resulting in the stress changing the variables in opposite directions. Values are the average of four technical replicates on at least 3 biological replicates of each genotype in each condition. The amount is expressed as percentage in weight of the cell wall. Boxes represent the mid two quartiles with the median drawn. The color represents the treatment. Control (blue) receiving a full Hoagland solution. Minus N (yellow) treated with a Hoagland solution without nitrogen. Minus P (orange) treated with a Hoagland solution lacking phosphorus. Different letters represent statistically significant values as determined by the Type III ANOVA and the Tukey HSD test (α = 0.05). Top, plot of the differences between different treatments on the same genotype. Bottom, plot of the differences between genotype under the same nutritional condition.

6.3.9 Cell wall architecture

The ratios between cellulose and hemicellulose (Cel/Hem) and arabinose and xylose (Ara/Xyl) in the cell wall were used as parameters for detecting differences in cell wall architecture in the biomass under study (Figure 6-15).

The nutrient stress affected the Ara/Xyl in Galahad (ANOVA, $p = 0.0074$). Both under nitrogen and phosphorus limiting conditions, the ratio in this genotype was significantly lower than in the control condition (t-test, $p = 0.029$ and $p= 0.016$, respectively).

Taken together these results indicate that under nutrient limitation, the decoration of xylans was contributing in determining the increased recalcitrance of the cell wall in Galahad, while it had no role in the decreased recalcitrance of Gaheris.

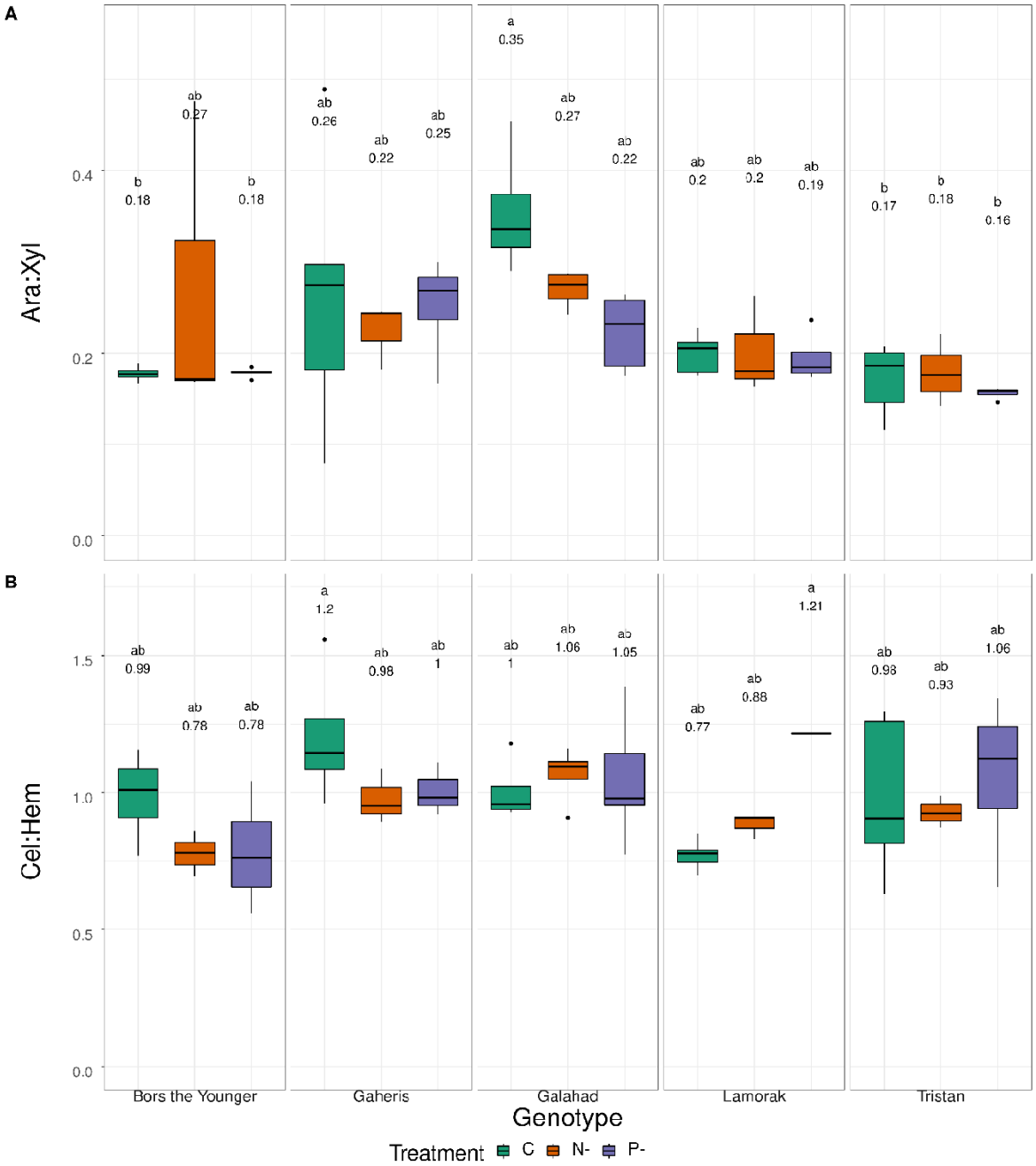


Figure 6-15 Effect of the treatment and genotype on the cellulose to hemicellulose and arabinose to xylose ratios.

The ratio between components of the cell wall have been found to strongly associate with the recalcitrance of the cell wall. Here we tested the effect of the nutrient stresses imposed (nitrogen limitation and phosphorous limitation) on the ratio between xylose and arabinose (A) and between cellulose and hemicellulose (B). There was a significant effect of the interaction between genotype and treatment. Values are the average of four technical replicates on at least 3 biological replicates of each genotype in each condition. The amount is expressed as percentage in weight of the cell wall. Boxes represent the mid two quartiles with the median drawn. The color represent the treatment. Control (blue) receiving a full Hoagland solution. Minus N (yellow) treated with a Hoagland solution without nitrogen. Minus P (orange) treated with a Hoagland solution lacking phosphorus. Different letters represent statistically significant values as determined by the Type III ANOVA and the Tukey HSD test ($\alpha = 0.05$). Top, plot of the differences between different treatments on the same genotype. Bottom, plot of the differences between genotype under the same nutritional condition.

6.4 Discussion

6.4.1 Plant architecture

A reduction in the amount of biomass produced was observed in the genotypes under nitrogen limitation, but not under phosphorus limitation. It was observed that the limitation could be mostly attributed to a lower tillering rather than a decrease in the weight of the single tiller. Indeed, the different allocation of biomass in several species of perennial grasses has been reported (Vazquez De Aldana et al., 1997). The observation of lower tillering in the genotypes of *Miscanthus* in this experiment could explain a reduction of rhizome budding. The effect of fertilisation is well known as a factor increasing the biomass yield of *Miscanthus* (Cosentino et al., 2007; Davis et al., 2015). Lee et al. (2017) found that the increase in biomass production in *Miscanthus* x giganteus fertilised with nitrogen was directly correlated with the number of tillers per plant. Recently the attention has moved to the effect of nutrient limitation. The reduction of biomass production as a consequence of limited nitrogen was reported in Sorghum bicolor (Makino and Ueno, 2018). When grown without water limitation, the nutrient stress reduced the biomass production of *Miscanthus* x giganteus and *Miscanthus* sacchariflorus. At the same time, *Miscanthus* sinensis was more resilient to nutrient limitation in this regard (da Costa et al., 2018). These results seem to agree with what was found in the present study, where the genotype of *M. sacchariflorus* and two of the hybrids were affected by the nutrient limitation, while *M. sinensis* did not respond to the treatment. This study expands the knowledge on the effect of nutrient limitation in *Miscanthus* since it was observed that also M.lutarioriparius and two genotypes of M.floridulus were similarly sensible to nutrient limitation.

The nutrient limitation did not have effect on the amount of biomass produced by the natural hybrid in this experiment. This is in contrast with previous literature where an

effect of nutrient limitation is often reported (da Costa et al., 2018). However, this could be explained by the fact that the natural hybrid used in this study was a different accession from ABR33. In da Costa et al. (2018) there is no evidence that the reduction in biomass production was related to the number of tillers or the weight of the single tiller. It is known that nutrient limitation can affect the biomass allocation in plants in a species-specific way (Yan et al., 2016). For example, species dominant under limited nitrogen environment tend to allocate more to leaves than stems (Yan et al., 2016). The reduction of tillers related to nitrogen limitation was reported in rice (Xiong et al., 2018) and related to phosphorus limitation in wheat (Rodríguez et al., 1998). However, there is no available evidence of mechanisms of competition between the number of tillers and weight of the tiller in other monocotyledons.

The effect of nutrient limitation on the leaf area was genotype-specific. Nitrogen stress determined a dramatic decrease of LA, while under phosphorus limitation, the trait was not significantly different from the one in the control condition. A reduction in leaf area under nitrogen limitation has been shown in maize (Massignam et al., 2011). In the study, the effect was more marked when the plants were grown in nitrogen limitation from the beginning. Also, phosphorus limitation has been seen to slow the development of new leaves in maize, without affecting the final size of the leaf (Usuda, 1995). Although the decrease of leaf area in *Miscanthus* was reported under drought stress, there is no evidence of the effect of nutrient limitation on area leaf in species of the genus. It is important here to remember that, differently from Maize, *Miscanthus* is a perennial grass using a different strategy to adapt to stressful conditions.

6.4.2 Plant physiology

6.4.2.1 Chlorophyll fluorescence

The chlorophyll fluorescence has been used as a method for assessing the of nutrient stress on photosynthesis in different species of plants (Kalaji et al., 2014a; Živčák et al., 2014b; Singh et al., 2019).

Here the nutrient stress affected the parameters of chlorophyll fluorescence in genotypes belonging to different species of *Miscanthus* with a strong effect of nitrogen and a moderate effect of phosphorus limitation on the same parameters.

Phosphorus limitation has been reported to affect the photosynthesis by inhibition of elements of the Photosystem I resulting on a flattening of step I of the JIP curve (Frydenvang et al., 2015; Carstensen et al., 2018b). In the present study, there was no effect of phosphorus limitation on the shape of the OJIP curve, and on the photosynthetic parameters derived. This result is in agreement with the one by Kalaji et al. (2014) on maize. In that study the authors tested the effect of various nutrient stresses on several parameters of the chlorophyll fluorescence curve in maize and found that while nitrogen limitation stress was decreasing the PITOT dramatically compared to the control, the phosphorus limitation stress was slightly increasing the parameter.

In the present study, an effect of nitrogen limitation on the ratio between Fv and Fm was observed in some of the genotypes under study under limiting nitrogen condition. There is no report of the effect of nitrogen limitation of photosynthetic parameters of *Miscanthus*, although the results presented here agree with what is reported in the literature for other plants. Similar behaviour was observed in maize, not fertilized (Wu et al., 2019) when compared with a fertilised condition. It is therefore possible that in the wild, germplasm with different tolerance to nutrient stress are available. This

agrees with the variability in soil condition where the genotypes were originally collected reported in Chapter 2.

6.4.2.2 Chlorophyll content

The chlorophyll content is related to the nutrient status of the plant. Here the nutrient stress affected the amount of chlorophyll in the 12 genotypes in the earliest time points of the experiment. The amount of chlorophyll remained to a low level under nitrogen limiting conditions. At the same time, it tended to increase to reach the same level of the control under phosphorus limitation. The effect of nitrogen on chlorophyll content has been known for long and relates to nitrogen being an integral component of the chlorophyll molecule (Schertz F. M., 1929; Tam and Magistad, 1935).

This is the first report of the effect of nitrogen and phosphorus limitation on chlorophyll content in *Miscanthus*. Results are in agreement with what has been observed in other plants.

In maize nitrogen limitation resulted in a lower content of chlorophyll with a different extent in two genotypes with different low -nitrogen tolerance (Wu et al., 2019). In maize (Usuda, 1995), low-P treatment induced the net loss of Chl in old leaf blades and phosphorus has been reported to limit the chlorophyll content in rice (Xu et al., 2007). The authors suggested that the phosphorus was acting on the photosynthesis through a limitation of the ATP supply.

6.4.3 Biomass quality

6.4.3.1 Saccharification

It has been forecasted that marginal lands are the best option for growing biomass crops. They are often affected by several stressful conditions. Nutrient limitation is one of these. We were her interested in determining if the stressful condition can alter the quality of biomass for conversion.

The results of this experiment revealed that nutrient limitation could affect the recalcitrance of biomass to conversion in genotypes of *Miscanthus*. Tang et al. (2018) found that, in Sorghum, low levels of nitrogen fertilization were associated with low levels of theoretical ethanol yields. In *Miscanthus*, da Costa et al. (2018) found that nutrient limitation was increasing monosaccharides enzymatically released in M. x giganteus and *M. sacchariflorus*, while there was no difference between treatment and control in *M. sinensis*. In the present study, we observed that the genotype of *M. sacchariflorus* (Gaheris) under nutrient limitation increased the release of monosaccharides, this seems to agree with what observed by da Costa et al. (2018). Conversely, it was here observed that one of the artificial hybrids (Galahad), under nitrogen limitation, had a lower enzymatic monosaccharide release, which is in contrast with what observed in natural hybrids by da Costa et al. (2018).

The nutritional stress induced a variation in the saccharification efficiency that changed the reciprocal ranking of the genotypes. This result could have a practical impact on the breeding of new *Miscanthus* varieties for the production of bio-based chemicals. The evaluation of the quality of the biomass obtained has to take into account alterations caused by specific nutritional conditions of the marginal lands where the crop is grown.

6.4.3.2 Cell wall composition and architecture

It was then investigated if nutrient status was able to induce a modification of the monosaccharide composition of the cell wall of the genotypes under study.

Firstly, the total amount of monosaccharide content was observed. The effect of nutrient stress on the total amount of monosaccharide content was specific according to the genotype considered and the mineral nutrient considered. In Galahad, the artificial hybrid with reduced saccharification under nutrient limitation, the nitrogen

limitation resulted in a significantly lower amount of glucose and xylose, but not in the amount of crystalline cellulose. Together with the observation that in this genotype the nutrient stress-induced a reduction of the arabinose to xylose ratio, this may suggest that the nutrient stress affected its biomass recalcitrance through an effect on its hemicellulose structure and composition. The result on the crystalline is in contrast with what found by da Costa et al. (2018) where all the three genotypes (M. sacchariflosus, *M. sinensis* and *M. x giganteus*) nutrient limitation had a lower content of crystalline cellulose.

Endotransglucosylases/hydrolases (XTH) are a class of enzymes acting on hemicellulose and involved in the remodelling of the cell wall under abiotic stress (Tenhaken, 2014). Most of the information on the effect of abiotic stress on these enzymes relates to experiment on heat and drought (Iurlaro et al., 2016). The information about the effect of nutrient stress on *Miscanthus* hemicellulose is limited. In Gaheris, the *M. sacchariflorus* that had increased saccharification under nutrient limitation, the phosphorus limitation resulted in a reduced content of monosaccharides. However, neither under nitrogen limitation, there was a difference in the monosaccharide composition and architecture of the cell wall. This may point to the fact that other components of the cell wall may have had a role in increasing the saccharification under nutrient limitation.

A more detailed test of cell wall composition and detailed architecture of the cell wall should be conducted on Galahad and Gaheris to elucidate the nature of the differences in the cell wall structure. These may include, but not limited to, methylation analysis, lignin content and composition, ferulic, p-coumaric and galacturonic acid determinations, immunolabelling and gene expression under nutrient stress.

7 Discussion

7.1 General discussion

Human history and the development of our society are strictly connected to the history of *Poales* (grasses) and their domestication (DeWet, 1981; Piperno et al., 2007). By providing staple food for the vast majority of humanity and food for the grazing animals from which humanity derives most of its protein, they shaped the cultural history of humans. Our species appeared on earth 100,000 years ago while grasses evolved in the Late Cretaceous (c. −80 Mya) (Prasad, 2005). Then, thousands of years ago, the relationship between the two species began with humans starting to grow grasses. That was the step that marked the transition of humans from hunter-gatherers to farmers. In present time, humanity is turning again to grasses to solve one of the biggest challenges of its history, seeking independence from fossil sources of energy. Using species of the genus *Miscanthus* as a source of biomass for the future biorefineries requires its domestication in the next decade (Yan et al., 2012; Robson et al., 2013). Studying *Miscanthus* poses several challenges compared to model plants and domesticated plants. The high level of heterozygosity, a consequence of the self-incompatibility in most species, the possibility of interspecific crosses, and the adaptation to different environments result in low replicability of experiments using this plant. Replicability can be challenging even in very standardised conditions, using clonal propagation. However, the study of non-model organisms is essential for creating knowledge that can help humanity in the face of climate change and energy shortage. It has been pointed out that the use of model plants has hindered our ability to understand stress resistance mechanisms. Model plants lack some of the mechanisms used by non-model plants (Voesenek et al., 2014).

Here a study of differences in terms of biomass quality between genotypes belonging to different genetic groups and collected from different geo-climatic areas was attempted. Some of the genotypes selected were exposed to nutrient stress in the attempt to understand the effect of abiotic stress on biomass quality. The main results have been discussed in the respective chapter. Here the key findings and the limitations of the research presented above are considered.

7.2 Key findings

7.2.1 Germplasm collection in the wild

The collection of *Miscanthus* germplasm in the wild in East Asia, in the last few decades, has resulted in an extensive germplasm collection in Europe (Huang et al., 2018) maintained at IBERS. Most of the collected variability was included in the ABR33 field trial. This germplasm is available to breeders for creating new varieties. However, the data suggest the existence of un-collected germplasm in several areas in East Asia with several species under-represented in ABR33. The collection and study of new wild accessions could provide access to variability for breeding programs.

The soil and climatic conditions in the areas where each species of *Miscanthus* is native were investigated. Results showed that the range of climatic and soil conditions in which species of the genus can be found is variable. Some of the environmental variates considered were describing the extent of environmental fluctuation of climate or soil conditions in each location. The 49 genotypes selected were potentially having different plasticity to different ecological conditions.

7.2.2 Effects of genotype and genetic group on biomass quality

The process of selecting *Miscanthus* varieties for biomass from wild accessions will inevitably result in the loss of genetic variance. Selective pressure on domestication genes in monocotyledons eliminated pre-existing variations in neighbouring regions

in a window spanning 10–100 kb (Glémin and Bataillon, 2009; Purugganan and Fuller, 2009). In *Zea mays* this generated the selective sweep at the gene Y1 locus producing the yellow kernel phenotype (Palaisa et al., 2004) and in rice the sweep in the gene waxy, which is associated with low-amylose rice in northeast Asian cultivars (Olsen and Purugganan, 2002)

The undesired loss of variability in loci not concerned with the domestication of *Miscanthus* could be lowered by the use of genetic selection. Particular attention should be put in preventing that the selection for high yields compromises the sustainability of biomass production (Robson et al., 2013). Therefore, the choice of the traits to target is crucial. Here it was shown that nutrient limitation can increase the amount of the monosaccharides that can be released enzymatically. Future breeding programs should take this into account by considering the nutrient use efficiency of the genotypes used and test the ability to maintain monosaccharide yield as well as biomass production under nutrient limiting conditions. This will reduce development of new varieties of *Miscanthus* more productive, but less resilient to abiotic stress. The results of the present study provided information about suitable target traits for biomass quality for conversion. The plant architecture and the ratio between leaf and stem are playing a crucial role in determining the quality of *Miscanthus* biomass for conversion (subsection 3.4.4). Here a difference in biomass quality for conversion between the material coming from leaf and material from the stem was shown. Also, variability in the leaf to stem ratio was found between the genetic groups included in this study (subsection 0). Due to the variation found, it is possible to say that the leaf to stem ratio could be a potential target for selection in *Miscanthus*. The ratio should be adapted according to the type of biomass desired and in the case where a material with low recalcitrance is required the preference should be given to leaf material, while

for conversion through combustion of thermochemical conversion (e.g. pyrolysis) more recalcitrant energy rich material from stems should be preferred. Considering the differences between leaf and stem biomass, genotypes from groups from *M. sacchariflorus* and *M.lutarioriparius* should be preferred where stem material is desired. At the same time, for applications requiring leaf characteristics, leafy genotypes such as *M. floridulus,* with a higher leaf to stem ratio, should be preferred. An attractive option would be provided by the interspecific crossing between the two groups. However, this possibility is still hindered by the differences in flowering time between species.

Another essential trait that should be taken into account in the process of selecting *Miscanthus* as a crop is the cell wall composition. New information on the extent of differences in cell wall composition in *Miscanthus* was found in the present study. The study provided information on cell wall composition in species of the genus *Miscanthus* that have been less investigated such as *Miscanthus floridulus* and *Miscanthus lutarioriparius*. The most important points to take in consideration are the effect of organ and the effect of nutrient limitation on the recalcitrance of the cell wall material and its architecture. Harnessing the mechanisms of genetic control of the link between plant architecture, cell wall structure and yield of chemical building blocks (glucans, monolignols, etc.) could dramatically increase the efficiency of the biorefining process. Robson et al. (2013) suggested that breeding programs could benefit from including *M. lutarioriparius*. Stavridou et al. (2019) found that *Miscanthus floridulus* could outperform M. x giganteus under conditions of no stress as well as single and multiple stress environments. This suggested that those genotypes would represent valuable targets for inclusion in future breeding programs. However, up to this moment, a detailed characterisation of the cell wall in genotypes

of *Miscanthus* other that *M. sinensis*, *M. sacchariflorus* and *M.* x *giganteus* has rarely been reported. In the conclusion of their study, Stavridou et al. (2019) suggested that their evidence could be the beginning point for mechanistic studies of stress in *Miscanthus*. Up to this moment, there is no evidence of investigations of nutrient stress on cell wall structure and composition in *Miscanthus*.

In this study, it was shown that leaf to stem differences in saccharification efficiency are underpinned by species-specific differences in cell wall composition. Although the difference in the total amount of hemicellulose between leaf and stem was not significant, the amount of galactose and arabinose was different between leaf and stem. Understanding the physiological and molecular basis of these differences could result in the identification of specific targets for the selection of genotypes with biomass suitable for specific conversion routes. It may be hypothesized that these targets will be found in promoter regions of genes involved in cell wall architecture and plant architecture. It is also possible to envision that promoter regions responding to abiotic stress and controlling the relationship between cell wall architecture and plant architecture could be identified and used as traits for the selection of new varieties.

Finally, the process of domestication of *Miscanthus* cannot forget to target the resilience to abiotic stress. The trait is particularly important for two main reasons. Firstly, climate change scenarios are predicted to produce an increased frequency in extreme weather events resulting in sudden alterations of the growth environment (e.g., flooding, rapid erosion). Secondly, the biomass crops, to be sustainable, will be grown on marginal lands, which will expose plants to abiotic stresses. At the same, to maintain the profitability of producing energy and chemicals from plants, biomass crops must be able to maintain the quality of their biomass under abiotic stress

conditions. In this context, it was here decided to investigate the effect of nutrient stress on traits related to biomass quality in different species of *Miscanthus* and their hybrids. Nutrient limitation is likely to happen on marginal lands as a consequence of the erosion of the topsoil.

7.2.3 Effects of nutrient limitation on biomass quality

The impact of abiotic stress on cell wall quality for conversion and cell wall composition was already documented. Here the knowledge was expanded by imposing nutrient limitation. Effects of nutrient limitation on plant cell walls have never been studied before in species of the genus *Miscanthus*. The information obtained by the experiments in the present work shows that there is variance in the reaction of *Miscanthus* species and their hybrids to nutrient stress in terms of adjustments to their cell wall properties and biomass quality (section 6.4.3.1). The response to nutrient limitation was different according to the genotype considered and the specific limitation imposed.

Interestingly, in this study, the two hybrids, coming from the same parental cross, had different shifts of biomass quality under nutrient stress conditions. Only Tristan under phosphorus limiting conditions released a significantly lower amount of the three main monosaccharides. The differences between the two hybrids were not mirrored by any of the cell wall characteristics measured (matrix monosaccharides, cellulose). The difference could reside in the quality of interlinking between the components in the cell wall such as the xylan decoration. Here we observed that the ratio between arabinose and xylose was only affected by nutrient stress in Galahad. Indeed it has been seen with xylem defective mutants of *Arabidopsis thaliana,* that the reduced functionality of the xylem and the dwarfism of the plants were associated with reduced acetylation of the xylans in the cell wall (Ramírez et al., 2018). This points to the need

for further investigation of the system interconnecting the synthesis and decoration of xylans in monocotyledons and how these are affected by abiotic stresses.

7.2.4 Evidence of plasticity of biomass quality traits in *Miscanthus*

Plasticity is a complex trait that was first identified in the '60s (Bradshaw, 1965). Plasticity refers to the extent of change of a physiologic or morphologic feature between environments. On the one hand, the ability of a genotype to maintain its characteristics when the environment changes are seen as a sign of adaptation and genetic stability, on the other hand, the ability to change as the environment changes (plasticity) can also be interpreted as a sign of adaptation. These contrasting observations are the result of the fact that plasticity itself is a trait under genetic control, and that responds to the same selection from the environment. Plasticity has been defined as the extent of environmental-related variability of a genotype. Since all the variability that is not genetic is environmental-related, it can be assumed that the concept of plasticity can be applied to all the intraspecific variation.

Plasticity can have two manifestations, morphological and physiological. All morphological changes have a physiological basis, so in principle, all the plasticity is physiological. However, the morphological plasticity relates to the physiological changes that cause a stable change in the morphology. The two types of plasticity are often a reciprocal relation, and morphological stability is underpinned by physiologic stability. For example, it was observed that the physiological mechanisms involved in the plant response to the form of nitrogen taken up, can result in cell wall remodelling with stable changes in cell wall thickness and composition (Podgórska et al. 2017).

If the duration of the environmental fluctuation is shorter than the generation time, any adaptation that occurs can only take place by plasticity (Bradshaw, 1965). In perennial grasses, the different range of conditions and the variability of the

environmental conditions in each environment to which each species is adapted can indicate the extent of its plasticity. For example, seeds of *Poa secunda* from more extreme, or stressful, climates (higher heat: moisture indices) tended to produce plants more plastic for panicle length and survival (Peirson, 2015). The senescence and yearly storage of nutrients in the rhizome is a typical example of plasticity in perennial plants. The loss of aboveground material is necessary for the adaptation of the plant to seasonal environmental variation. At the same time, the translocation of nutrients minimises the deleterious effects of organ loss (Bradshaw, 1965). Conversely, plasticity is less likely in annual plants, where plasticity does not confer any advantage if it is not easily reversible.

Plasticity is a property specific to individual characters concerning particular environmental conditions. The results of the present study in terms of the effect of genotype, organ on plant biomass characteristics under normal conditions and nutrient limiting condition could be interpreted as evidence that the extent of plasticity in the genus *Miscanthus* is variable between species, possibly as a consequence of the adaptation of the species to different pedo-climatic environments. In the present study, it was observed that under nutrient stress, the number of tillers rather than the weight of the single tiller was affected. This could be interpreted as a consequence of the perennial habit of *Miscanthus* and its resource management strategy. Indeed, it has been seen that nutrient limitation in soil can influence biomass allocation in several species of perennial grasses (Yan et al., 2016). .

Future experiments on the effect of nutrient stress in *Miscanthus* should also aim include the quantification of this trait and its genetic basis. This knowledge could help breeding programs whose goal is the creation of varieties for specific marginal lands.

The alterations caused by specific nutritional conditions of the marginal lands where the crop is grown on the quality of the biomass obtained has to be considered during the choice of the genotype. The aim could either be predicting the quality under nutrient stress conditions or use specific fertilisation plans to alter biomass quality and tailor it to biorefinery needs.

7.2.5 *Miscanthus* biomass for biorefineries: reconciling quantity and quality

A future scenario where bio-refineries will replace oil-based refineries for the production of fuel and chemicals and *Miscanthus* grown on marginal lands is going to be one of the primary sources for their feedstock can be envisioned. Similarly to what is happening for sugarcane and pulpwood in South American and oil palm in Southeast Asia, it is possible to envisage a business model that integrates biorefineries and feedstock plantations (Dou et al., 2018). The key for the system to be profitable and sustainable is to find a compromise between crop production (growers profit) and conversion quality of the biomass (biorefinery profit). One possible approach is that biorefineries could pay for the biomass by quality in contrast with the payment based on the dry matter. Dou et al. (2018) studied the advantages related to the use of two hybrids of poplar for biorefineries and farming them on different lands. They found that when quality for the conversion is taken into account, the situation is inverted and suggested the use of an overarching variable considering yield and conversion quality to assess the productivity of a hybrid. It is foreseeable that farmers could require in turn varieties of *Miscanthus* more productive and with specific biomass traits, such as lower recalcitrance to enzymatic deconstruction, higher yield of building block for biorefining and stability of the traits under stress conditions.

In the near future, one of the main goals of the research will be to provide the growers with tools to achieve better biomass quality. These tools will be, on the one hand, new

germplasm, able to give a profitable yield of good quality biomass on marginal lands and, on the other hand, appropriate agronomic advice to increase the quality of the biomass. Advice has to include land management, fertilisation and crop protection tailored to the need for traits required for specific biorefining activities. Improving the knowledge of the effect of single and multiple abiotic stresses will contribute to developing effective strategies of crop management. The ability to select new germplasm could benefit from the knowledge of specific geo-climatic areas where genotypes with superior quality can be found.

7.3 Limitations

The experimental work presented here was limited in some aspects. One of the main limitations of the experiments in this work was the rhizomatous perennial habit of *Miscanthus*. The perenniality in *Miscanthus* is associated with a juvenile phase that lasts for 2 – 5 years. During these years, there is an increase in yields, but also a change in the cell wall composition (van der Weijde et al., 2017a). Here the genotypes were grown under limiting nutrient conditions for 60 days after emergence. While the results can be used to compare relative differences, they cannot be used to infer the effects of nutrient limitation on mature plants. The relation between the impact of stress in an early stage of a plant and the quality of the biomass at a later stage has never been tested.

The biomass of the genotypes grown in a field in Wales was compared and used to describe the quality of the biomass of each genotype for conversion. Differences in climatic conditions between the area of collection of the biomass and the conditions where each genotype was originally collected were estimated. However, it was impossible to collect plant material from the plants of the same genotypes grown under natural conditions in the environment where they are adapted. The comparison

between the two biomass samples could provide evidence of the ability of each genotype to perceive the different environmental conditions.

Plants in the experiment under nutrient stress condition were exposed for 60 days to the low content of nitrogen or phosphorus in the soil. The treatment resulted in effects on biomass composition. However, it has to be pointed out that the level of stress imposed is unlikely to happen in field conditions. The decision of imposing this level of stress was taken consciously considering different factors. First of all, the nutrient limitation was here tested not only as a source of stress but as a possible treatment to generate specific alterations of the biomass quality. It cannot be excluded that a moderate reduction of the level of the nutrient in the soil would not have caused any symptoms or effects on biomass properties. The long cycle of *Miscanthus* and the limited amount of time did not allow for repeating the experiment using more levels of each nutrient. Finally, it is essential to remember that there was a pool of nutrients stored in the rhizome. In field conditions, where a nutrient is lacking, it is possible to assume that also the level of nutrients stored in the rhizome will be low. Thus, the deficient level of nutrients was aimed to avoid that the plant could prevent the stress on the short period by using its rhizome reserves. Besides, the measurements of the photosynthetic parameters showed a moderate level of stress perceived. It is advisable that experiments performed to determine nutrient use efficiency and nutrient stress in plants follow a series of preliminary experiments aimed to characterise the physiology of the plants under non-stressful conditions using small randomised plots and standardised conditions.

Differences between leaf and stem in terms of biomass composition have been reported in *Miscanthus*. Here we found that the differences between leaf and stem in terms of cell wall composition were genotype-related. It could not be excluded that

alterations of the biomass induced by the nutrient stress could impact on the organ-related differences. However, at the moment of biomass harvesting, most of the genotype had not developed a stem, and it was impossible to compare samples from different organs.

Molecular biology techniques could have helped in confirming some of the hypothesis formulated in this thesis. However, the lack of a complete sequence of the *Miscanthus* genome and a limited characterisation of genes limited the ability to perform such tests.

The cell wall composition was tested in terms of monosaccharide content in the matrix, cellulose content, lignin content and composition. Although this approach provided useful information, especially in the case of the nutrient stress experiment, the characterization presents several gaps. It was noticed that the nutrient stress did not affect the total amount of hemicellulose, but possibly the xylan decoration. However, this was assessed only using the ratio between arabinose and xylose. More detailed information could be obtained by esterification analysis (Pettolino et al., 2012; Petit et al., 2019), *p*-coumaric (*p*-CA) and ferulic (FA) acid contents, and *in vivo* immunolabelling. All these characteristics have been indicated as predictive of the quality of cell wall for conversion (da Costa et al., 2019)

The limitations summarised here were crucial to determining the future directions of the research. At the same time, the principal results pointed to the new research questions that are important to answer in the near future.

7.4 Future directions

7.4.1 The need for further investigation

The experimental work presented here pointed to the need for more evidence on some points. The quantification of plant plasticity under different stress conditions is an

important trait. It could help the selection of novel germplasm to introduce in breeding programs. Further investigation using more genotypes/species and under different nutrition conditions could contribute to quantifying more accurately the extent of the plasticity of *Miscanthus* biomass quality in different nutrient scenarios.

Further studies of the response of *Miscanthus* species to nutrient limitation should include testing for more nutrient conditions, such as deprivation of micro and micro-nutrients and different level of reduction of single and multiple nutrients, restoration experiments in which the nutrient is provided again partially or totally after a deprivation period. To better understand the correlation between phenologic status and stress response in perennial plants, it would be interesting to perform experiments in which the stress condition is imposed to plants at different growth stages.

The preliminary information created in this thesis could be completed with molecular tests aimed to investigate the mechanism of connection between nutrient stress and phenological/histological alterations. The study could involve the investigation of nutrient uptake mechanisms, nutrient translocation, nutrient remobilization and storage in rhizomes, nutrient metabolization.

A more detailed test of cell wall composition and fine architecture of the cell wall should be conducted to elucidate the nature of the differences in the cell wall structure. Especially it would be interesting performing the tests on the two sibling genotypes Galahad and Gaheris. These may include, but not limited to, methylation analysis, lignin content and composition, ferulic, p-coumaric and galacturonic acid determinations, gene expression under nutrient stress.

7.4.2 New research questions

Miscanthus is a fascinating plant, and its study could help to better understand many characteristics related to perenniality and the production of biomass for biorefineries.

From the results obtained here, new questions arose. First of all, the influence of nutrient limitation on perennial plants with C4 metabolism is different from annual plants. Its understanding could help not only the improvement of *Miscanthus* but also inform the development of other monocotyledon crops, both perennial and annual.

It is particularly interesting to determine the relationship between juvenile and mature stages in perennial plants. The definition of these stages is still controversial, and the understanding of the effect of the environment yet not explored. Questions that should be answered are “What is the effect of environmental conditions experienced in one stage on the other stages of the plant?”, “What is the function of the rhizome under abiotic stress conditions”, “What is the effect of abiotic stress on the reallocation of resources”, “Can stress perceived during dormancy affecting biomass quality in the next season”.

The answer to these and other related questions would not only be interesting from a scientific point of view. It could also impact on the ability to use marginal lands to produce energy and chemicals from plants.

References

Acquah GE, Via BK, Fasina OO, Adhikari S, Billor N, Eckhardt LG (2017) Chemometric modeling of thermogravimetric data for the compositional analysis of forest biomass. PLoS One **12**: 1–15

Adams JMM, Winters AL, Hodgson EM, Gallagher JA (2018) What cell wall components are the best indicators for *Miscanthus* digestibility and conversion to ethanol following variable pretreatments? Biotechnol Biofuels **11**: 1–14

Albersheim P, Darvil A, Roberts K, Sederof R, Staehelin A (2011) Plant Cell Walls: From Chemistry to Biology. Q Rev Biol. doi: 10.1086/662480

Albersheim P, Nevins DJ, English PD, Karr A (1967) A method for the analysis of sugars in plant cell-wall polysaccharides by gas-liquid chromatography. Carbohydr Res **5**: 340–345

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). The Plant Cell Wall. In Molecular Biology of the Cell. 4th edition. Garland Science. https://www.ncbi.nlm.nih.gov/books/NBK26928/

Allison GG, Morris C, Clifton-Brown J, Lister SJ, Donnison IS (2011). Genotypic variation in cell wall composition in a diverse set of 244 accessions of *Miscanthus*. Biomass and Bioenergy, 35(11), 4740–4747. https://doi.org/10.1016/j.biombioe.2011.10.008

Anitha Gopal B, Muralikrishna G (2009) Porcine pancreatic -amylase and its isoforms: Purification and kinetic studies. Int J Food Prop **12**: 571–586

Antal MJ (1983) Biomass Pyrolysis: A Review of the Literature Part 1—Carbohydrate Pyrolysis BT - Advances in Solar Energy: An Annual Review of Research and Development, Volume 1 · 1982. *In* KW Böer, JA Duffie, eds,Springer New York, Boston, MA, pp 61–111

Arnoult S, Brancourt-Hulmel M (2014). A Review on *Miscanthus* Biomass Production and Composition for Bioenergy Use: Genotypic and Environmental Variability and Implications for Breeding. BioEnergy Research, 8(2), 502–526. https://doi.org/10.1007/s12155-014-9524-7

Arrhenius S (1897) On the Influence of Carbonic Acid in the Air upon the Temperature of the Earth. Publ Astron Soc Pacific **9**: 14

Aspinwall MJ, Loik ME, Resco de Dios V, Tjoelker MG, Payton PR, Tissue DT (2015) Utilizing intraspecific variation in phenotypic plasticity to bolster agricultural and forest productivity under climate change. Plant Cell Environ **38**: 1752–1764

Assuero SG, Mollier A, Pellerin S (2004) The decrease in growth of phosphorus-deficient maize leaves is related to a lower cell production. Plant, Cell Environ **27**: 887–895

Auguie B (2017) gridExtra: Miscellaneous Functions for "Grid" Graphics.

Barker A V, Pilbeam DJ (2016) Handbook of Plant Nutrition. CRC Press

Bernardo J (2014) Biologically grounded predictions of species resistance and resilience to climate change. Proc Natl Acad Sci U S A **111**: 5450–5451

Bhatia R, Dalton S, Roberts LA, Moron-Garcia OM, Iacono R, Kosik O, Gallagher JA, Bosch M (2019) Modified expression of ZmMYB167 in Brachypodium distachyon and Zea mays leads to increased cell wall lignin and

phenolic content. Sci Rep. doi: 10.1038/s41598-019-45225-9
Bhatia R, Winters A, Bryant DN, Bosch M, Clifton-Brown J, Leak D, Gallagher J (2020). Pilot-scale production of xylo-oligosaccharides and fermentable sugars from *Miscanthus* using steam explosion pretreatment. Bioresource Technology, **296**(September 2019), 122285. https://doi.org/10.1016/j.biortech.2019.122285
Bin Y, Hongzhang C (2010) Effect of the ash on enzymatic hydrolysis of steam-exploded rice straw. Bioresource Technology, 101(23), 9114–9119. https://doi.org/10.1016/j.biortech.2010.07.033
Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol, 54, 519–546. https://doi.org/10.1146/annurev.arplant.54.031902.134938
Bozell JJ, Petersen GR (2010) Technology development for the production of biobased products from biorefinery carbohydrates—the US Department of Energy's "Top 10" revisited. Green Chem **12**: 539–554
Bradshaw AD (1965) Evolutionary Significance of Phenotypic Plasticity in Plants. *In* EW Caspari, JMBT-A in G Thoday, eds,Academic Press, pp 115–155
Bro R, Kiers HAL (2003) A new efficient method for determining the number of components in PARAFAC models. J Chemom **17**: 274–286
Buchanan BB, Gruissem W, Jones RL, eds (2000) Biochemistry & molecular biology of plants. Rockville, Md. : American Society of Plant Physiologists, Rockville, Md.
Butler E, Devlin G, Meier D, McDonnell K (2013) Characterisation of spruce, salix, *Miscanthus* and wheat straw for pyrolysis applications. Bioresour Technol **131**: 202–209
Campbell MM, Sederoff RR (1996) Variation in lignin content and composition. Plant Physiol **110**: 3–13
Cao Y, Li J, Yu L, Chai G, He G, Hu R, Qi G, Kong Y, Fu C, Zhou G (2014) Cell wall polysaccharide distribution in *Miscanthus* lutarioriparius stem using immuno-detection. Plant Cell Rep **33**: 643–653
Carlsson G, Mårtensson LM, Prade T, Svensson SE, Jensen ES (2017) Perennial species mixtures for multifunctional production of biomass on marginal land. GCB Bioenergy **9**: 191–201
Carpita NC, Gibeaut DM (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. Plant J **3**: 30
Carpita NC, McCann MC (2015) Characterizing visible and invisible cell wall mutant phenotypes. J Exp Bot **66**: 4146–4163
Carstensen A, Herdean A, Schmidt SB, Sharma A, Spetea C, Pribil M, Husted S (2018a) The impacts of phosphorus deficiency on the photosynthetic electron transport chain. Plant Physiol **177**: pp.01624.2017
Carstensen A, Herdean A, Schmidt SB, Sharma A, Spetea C, Pribil M, Husted S (2018b) The impacts of phosphorus deficiency on the photosynthetic electron transport chain. Plant Physiol **177**: pp.01624.2017
Cass CL, Lavell AA, Santoro N, Foster CE, Karlen SD, Smith RA, Ralph J, Garvin DF, Sedbrook JC (2016) Cell wall composition and biomass recalcitrance differences within a genotypically diverse set of Brachypodium distachyon inbred lines. Front Plant Sci **7**: 1–16
Cattell RB (Duk. U (1944) "Parallel proportional profiles" and other principles for determining the choice of factor rotation. Psychometrika **9**: 267–283

Cerazy-Waliszewska J, Jeżowski S, Łysakowski P, Waliszewska B, Zborowska M, Sobańska K, Ślusarkiewicz-Jarzina A, Białas W, Pniewski T (2019) Potential of bioethanol production from biomass of various *Miscanthus* genotypes cultivated in three-year plantations in west-central Poland. Ind Crops Prod **141**: 111790

Cheng H, Wang L (2013) Lignocelluloses Feedstock Biorefinery as Petrorefinery Substitutes. Biomass Now - Sustain. Growth Use. InTech, p 13

Christensen U, Alonso-Simon A, Scheller H V., Willats WGT, Harholt J (2010) Characterization of the primary cell walls of seedlings of Brachypodium distachyon - A potential model plant for temperate grasses. Phytochemistry **71**: 62–69

Clark JH, Deswarte FEI (2008) The Biorefinery Concept–An Integrated Approach. Introd to Chem from Biomass 1–20

Clark JH, Luque R, Matharu AS (2012) Green Chemistry, Biofuels, and Biorefinery. Annu Rev Chem Biomol Eng **3**: 183–207

Clark L V, Brummer JE, Glowacka K, Hall MC, Heo K, Peng J, Yamada T, Yoo JH, Yu CY, Zhao H, et al (2014) A footprint of past climate change on the diversity and population structure of *Miscanthus* sinensis. Ann Bot **114**: 97–107

Clark L V, Dzyubenko E, Dzyubenko N, Bagmet L, Sabitov A, Chebukin P, Johnson DA, Kjeldsen JB, Petersen KK, Jørgensen U, et al (2016) Ecological characteristics and in situ genetic associations for yield-component traits of wild *Miscanthus* from eastern Russia. Ann Bot **118**: 941–955

Clayton WD, Vorontsova MS, Harman KT, Williamson H (2006) GrassBase - The Online World Grass Flora. http://www.kew.org/data/grasses-db.html

Clifton-Brown J, Lewandowski I, Andersson B, Basch G, Christian DG, Kjeldsen JB, Jørgensen U, Mortensen J V, Riche A, Schwarz KU, et al (2001) Performance of 15 *Miscanthus* Genotypes at Five Sites in Europe. Agron J **93**: 7

Clifton-Brown J C, Lewandowski I (2000). Water Use Efficiency and Biomass Partitioning of Three Different *Miscanthus* Genotypes with Limited and Unlimited Water Supply. Ann Bot, 86(1), 10. https://doi.org/10.1006/anbo.2000.1183

Cordell D, Drangert JO, White S (2009). The story of phosphorus: Global food security and food for thought. Global Environmental Change, 19(2), 292–305. https://doi.org/10.1016/j.gloenvcha.2008.10.009

Clifton-Brown J, Schwarz KU, Awty-Carroll D, Iurato A, Meyer H, Greef J, Gwyn J, Mos M, Ashman C, Hayes C, et al (2019) Breeding strategies to improve *Miscanthus* as a sustainable source of biomass for bioenergy and biorenewable products. Agronomy. doi: 10.3390/agronomy9110673

Cosentino SL, Patané C, Sanzone E, Copani V, Foti S (2007) Effects of soil water content and nitrogen supply on the productivity of *Miscanthus* x giganteus Greef et Deu. in a Mediterranean environment. Ind Crops Prod **25**: 75–88

Cosgrove DJ, Jarvis MC (2012). Comparative structure and biomechanics of plant primary and secondary cell walls. Frontiers in Plant Science, 3. https://doi.org/10.3389/fpls.2012.00204

da Costa RMF, Pattathil S, Avci U, Winters A, Hahn MG, Bosch M (2019) Desirable plant cell wall traits for higher-quality *Miscanthus* lignocellulosic biomass. Biotechnol Biofuels **12**: 1–18

da Costa RMF, Simister R, Roberts LA, Timms-Taravella E, Cambler AB,

Corke FMK, Han J, Ward RJ, Buckeridge MS, Gomez LD, et al (2018) Nutrient and drought stress: implications for phenology and biomass quality in *Miscanthus*. Ann Bot 1–14**da Costa RMF, Pattathil S, Avci U, Lee SJ, Hazen SP, Winters A, Hahn MG, Bosch M** (2017) A cell wall reference profile for *Miscanthus* bioenergy crops highlights compositional and structural variations associated with development and organ origin. New Phytol **213**: 1710–1725

Costa RMF da (2015) Constructing a Comprehensive Picture of *Miscanthus* Cell Wall to Advance its Deconstruction. Aberystwyth University

da Costa RMF, Lee SJ, Allison GG, Hazen SP, Winters A, Bosch M (2014) Genotype, development and tissue-derived variation of cell-wall properties in the lignocellulosic energy crop *Miscanthus*. Ann Bot **114**: 1265–1277

Dauber J, Brown C, Fernando AL, Finnan J, Krasuska E, Ponitka J, Styles D, Thrän D, Van Groenigen KJ, Weih M, et al (2012) Bioenergy from "surplus" land: Environmental and socio-economic implications. BioRisk **50**: 5–50

Davis MP, David MB, Voigt TB, Mitchell CA (2015) Effect of nitrogen addition on *Miscanthus* 3 giganteus yield, nitrogen losses, and soil organic matter across five sites. GCB Bioenergy **7**: 1222–1231

Davis R, Grundl N, Tao L, Biddy MJ, Tan ECD, Beckham GT, Humbird D, Thompson DN, Roni MS (2018) Process Design and Economics for the Conversion of Lignocellulosic Biomass to Hydrocarbon Fuels and Coproducts: 2018 Biochemical Design Case Update: Biochemical Deconstruction and Conversion of Biomass to Fuels and Products via Integrated Biorefinery Path.

Demirbas A (2005) Potential applications of renewable energy sources, biomass combustion problems in boiler power systems and combustion related environmental issues. Prog Energy Combust Sci **31**: 171–192

Deng Y, Parajuli P B (2016). Return of investment and profitability analysis of bio-fuels production using a modeling approach. Information Processing in Agriculture, **3**(2), 92–98. https://doi.org/10.1016/j.inpa.2016.03.002

De Souza AP, Kamei CLAA, Torres AF, Pattathil S, Hahn MG, Trindade LM, Buckeridge MS, Alvim Kamei CL, Torres AF, Pattathil S, et al (2015) How cell wall complexity influences saccharification efficiency in *Miscanthus* sinensis. J Exp Bot **66**: 4351–4365

DeWet JMJ (1981) Grasses and the Culture History of Man. Ann Missouri Bot Gard **68**: 87–104

Domon JM, Baldwin L, Acket S, Caudeville E, Arnoult S, Zub H, Gillet F, Lejeune-Henaut I, Brancourt-Hulmel M, Pelloux J, et al (2013) Cell wall compositional modifications of *Miscanthus* ecotypes in response to cold acclimation. Phytochemistry **85**: 51–61

Dou C, Gustafson R, Bura R (2018) Bridging the gap between feedstock growers and users: the study of a coppice poplar-based biorefinery. Biotechnol Biofuels **11**: 77

Elser JJ (2012). Phosphorus: A limiting nutrient for humanity? Current Opinion in Biotechnology, 23(6), 833–838. https://doi.org/10.1016/j.copbio.2012.03.001

Engels C, Kirkby E, White P (2011) Mineral Nutrition, Yield and Source-Sink Relationships. Marschner's Miner Nutr High Plants Third Ed. doi: 10.1016/B978-0-12-384905-2.00005-4

Fahmi R, Bridgwater A V, Donnison I, Yates N, Jones JM (2008) The effect of lignin and inorganic species in biomass on pyrolysis oil yields, quality and stability. Fuel **87**: 1230–1240

Fahmy TYA, Fahmy Y, Mobarak F, El-Sakhawy M, Abou-Zeid RE (2020) Biomass pyrolysis: past, present, and future. Environ Dev Sustain **22**: 17–32

Fangel JU, Ulvskov P, Knox JP, Mikkelsen MD, Harholt J, Popper ZA, Willats WGT (2012). Cell wall evolution and diversity. Frontiers in Plant Science, 3(July), 1–8. https://doi.org/10.3389/fpls.2012.00152

Fao/Iiasa/Isric/Isscas/Jrc (2009) Harmonized World Soil Database (version 1.1). FAO, Rome, Italy IIASA, Laxenburg, Austria

Fehér C (2018) Novel approaches for biotechnological production and application of L-arabinose. J Carbohydr Chem **37**: 251–284

Feist M (2015) Thermal analysis: basics, applications, and benefit. ChemTexts **1**: 8

Fisher RA (1919) XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. Trans R Soc Edinburgh **52**: 399–433

Foster CE, Martin TM, Pauly M (2010) Comprehensive Compositional Analysis of Plant Cell Walls (Lignocellulosic biomass) Part I: Lignin. J Vis Exp 5–8

Foston M (2014) Biomass recalcitrance and the contributing call wall factors. *In* AJ Ragauskas, ed, Mater. Biofuels. World Scientific Publishing Co Pte Ltd, Singapore, pp 29–44

Fry SC (1988) The growing plant cell wall: chemical and metabolic analysis. Longman Scientific & Technical

Frydenvang J, van Maarschalkerweerd M, Carstensen A, Mundus S, Schmidt SB, Pedas PR, Laursen KH, Schjoerring JK, Husted S (2015) Sensitive detection of phosphorus deficiency in plants using chlorophyll a fluorescence. Plant Physiol **169**: 353–361

Fukushima RS, Hatfield RD (2001) Extraction and isolation of lignin for utilization as a standard to determine lignin concentration using the acetyl bromide spectrophotometric method. J Agric Food Chem **49**: 3133–3139

Le Gall H, Philippe F, Domon J-M, Gillet F, Pelloux J, Rayon C (2015) Cell Wall Metabolism in Response to Abiotic Stress. Plants **4**: 112–166

Geldner N (2013) The endodermis. Annu Rev Plant Biol **64**: 531–558

Gelfand I, Sahajpal R, Zhang X, Izaurralde RC, Gross KL, Robertson GP (2013) Sustainable bioenergy production from marginal lands in the US Midwest. Nature **493**: 514–517

Gilbert N (2009) The Disappearing Nutrient. Nature **461**: 716–718

Głazowska S, Baldwin L, Mravec J, Bukh C, Hansen T H, Jensen M M, Fangel J U, Willats W G T, Glasius M, Felby C, Schjoerring J K (2018). The impact of silicon on cell wall composition and enzymatic saccharification of Brachypodium distachyon. Biotechnology for Biofuels, 11(1), 0–18. https://doi.org/10.1186/s13068-018-1166-0

Glémin S, Bataillon T (2009) A comparative view of the evolution of grasses under domestication. New Phytol **183**: 273–290

Gomez LD, Steele-King CG, McQueen-Mason SJ (2008) Sustainable liquid biofuels from biomass: the writing's on the walls. New Phytol **178**: 473–485

Gorban AN, Pokidysheva LI, Smirnova E V., Tyukina TA (2011) Law of the Minimum Paradoxes. Bull Math Biol **73**: 2013–2044

Grabber JH, Ralph J, Hatfield RD (1998) Modeling Lignification in Grasses with Monolignol Dehydropolymerisate-Cell Wall Complexes. Lignin Lignan Biosynth.

American Chemical Society, pp 12–163
Grantham NJ, Wurman-Rodrich J, Terrett OM, Lyczakowski JJ, Stott K, Iuga D, Simmons TJ, Durand-Tardif M, Brown SP, Dupree R, et al (2017) An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls. Nat Plants **3**: 859–865
Greenhalf CE, Nowakowski DJ, Harms AB, Titiloye JO, Bridgwater A V. (2013) A comparative study of straw, perennial grasses and hardwoods in terms of fast pyrolysis products. Fuel **108**: 216–230
Gruner EL (1875) Traité de métallurgie: 1: Agents et appareils métallurgiques. Principes de la combustion. Dunod
Guggisberg A, Liu X, Suter L, Mansion G, Fischer M C, Fior S, Roumet M, Kretzschmar R, Koch M A, Widmer A (2018) The genomic basis of adaptation to calcareous and siliceous soils in Arabidopsis lyrata. Molecular Ecology, **27**(24), 5088–5103. https://doi.org/10.1111/mec.14930
Hager HA, Sinasac SE, Gedalof Z, Newman JA (2014) Predicting Potential Global Distributions of Two *Miscanthus* Grasses : Implications for Horticulture , Biofuel Production , and Biological Invasions. PLoS One **9**: 1–14
Hahn MG, Darvill AG, Albersheim P (1981) Host-Pathogen Interactions. Plant Physiol **68**: 1161–1169
Harris PJ, Hartley RD (1976) Detection of bound ferulic acid in cell walls of the Gramineae by ultraviolet fluorescence microscopy. Nature **259**: 508–510
Hatfield RD, Grabber J, Ralph J, Brei K (1999) Using the acetyl bromide assay to determine lignin concentrations in herbaceous plants: Some cautionary notes. J Agric Food Chem **47**: 628–632
Hatfield RD, Marita JM, Frost K (2008) Characterization of p-coumarate accumulation, p-coumaroyl transferase, and cell wall changes during the development of corn stems. J Sci Food Agric **88**: 2529–2537
Hatfield RD, Rancour DM, Marita JM (2017) Grass Cell Walls: A Story of Cross-Linking. Front Plant Sci. doi: 10.3389/fpls.2016.02056
Havaux M, Strasser RJ, Greppin H (1991) A theoretical and experimental analysis of the qP and qN coefficients of chlorophyll fluorescence quenching and their relation to photochemical and nonphotochemical events. Photosynth Res **27**: 41–55
Hawkesford MJ, Horst W, Kichey T, Lambers H, Schjoerring, Jan Kofod, Skrumsager , Inge, White P (2011) Functions of Macronutrients. Marschner's Miner. Nutr. High. Plants. Elsevier Science & Technology, San Diego, UNITED KINGDOM, pp 135–190
Hegerl GC, Brönnimann S, Cowan T, Friedman AR, Hawkins E, Iles C, Müller W, Schurer A, Undorf S (2019) Causes of climate change over the historical record. Environ Res Lett. doi: 10.1088/1748-9326/ab4557
Held MA, Jiang N, Basu D, Showalter AM, Faik A (2014) Plant Cell Wall Polysaccharides: Structure and Biosynthesis. In K. G. Ramawat & J.-M. Mérillon (Eds.), Polysaccharides (pp. 1–47). Springer International Publishing. https://doi.org/10.1007/978-3-319-03751-6_73-1

Herr DG (1986) On the history of ANOVA in unbalanced, factorial designs: The first 30 years. Am Stat **40**: 265–270
Hijmans RJ, Elith J (2017) [R Manual] Species distribution modeling with R. Species Distrib Model with R 79

Himmel ME, Ding S, Johnson DK, Adney WS, Mark R, Brady JW, Foust TD (2016) Biomass Recalcitrance : Engineering Plants and Enzymes for Biofuels Production. Science (80-) **315**: 804–807

Hoagland DR, Arnon DI, Arnon Revised DI, Arnon DI (1950) The water-culture method for growing plants without soil. Calif Agric Exp Stn Circ **347**: 1–32

Hodgson EM, Fahmi R, Yates N, Barraclough T, Shield I, Allison G, Bridgwater A V, Donnison IS (2010) *Miscanthus* as a feedstock for fast-pyrolysis: does agronomic treatment affect quality? Bioresour Technol **101**: 6185–6191

Hodgson EM, Nowakowski DJ, Shield I, Riche A, Bridgwater A V., Clifton-Brown JC, Donnison IS (2011) Variation in *Miscanthus* chemical composition and implications for conversion by pyrolysis and thermo-chemical bio-refining for fuels and chemicals. Bioresour Technol **102**: 3411–3418

Hodkinson TR, Cesare M de, Barth S (2013) Nuclear SSR Markers for *Miscanthus*, Saccharum, and Related Grasses (Saccharinae, Poaceae) . Appl Plant Sci **1**: 1300042

Hodkinson TR, Chase MW, Lledo MD, Salamin N, Renvoize SA (2002) Phylogenetics of *Miscanthus*, Saccharum and related genera (Saccharinae, Andropogoneae, Poaceae) based on DNA sequences from ITS nuclear ribosomal DNA and plastid trnLintron and trnL-F intergenic spacers. J Plant Res **115**: 381–392

Hodkinson TR, Klaas M, Jones MB, Prickett R, Barth S (2015) *Miscanthus* : a case study for the utilization of natural genetic variation. Plant Genet Resour **13**: 219–237

Hodkinson TR, Renvoize S (2001) Nomenclature of *Miscanthus* xgiganteus (Poaceae). Kew Bull **56**: 759

Hoehenwarter W, Mönchgesang S, Neumann S, Majovsky P, Abel S, Müller J (2016) Comparative expression profiling reveals a role of the root apoplast in local phosphate response. BMC Plant Biol **16**: 1–21

Huang LS, Flavell R, Donnison IS, Chiang Y-C, Hastings A, Hayes C, Heidt C, Hong H, Hsu T-W, Humphreys M, et al (2018) Collecting wild *Miscanthus* germplasm in Asia for crop improvement and conservation in Europe whilst adhering to the guidelines of the United Nations' Convention on Biological Diversity . Ann Bot 1–14

Ikram S, Bedu M, Daniel-Vedele F, Chaillou S, Chardon F (2012) Natural variation of Arabidopsis response to nitrogen availability. J Exp Bot **63**: 91–105

Ings J, Mur LA, Robson PR, Bosch M (2013) Physiological and growth responses to water deficit in the bioenergy crop *Miscanthus* x giganteus. Front Plant Sci **4**: 468

Ishii T, Hiroi T (1990) Isolation and characterization of feruloylated arabinoxylan oligosaccharides from bamboo shoot cell-walls. Carbohydr Res **196**: 175–183

Iurlaro A, De Caroli M, Sabella E, De Pascali M, Rampino P, De Bellis L, Perrotta C, Dalessandro G, Piro G, Fry SC, et al (2016) Drought and heat differentially affect XTH expression and XET activity and action in 3-day-old seedlings of durum wheat cultivars with different stress susceptibility. Front Plant Sci **7**: 1–18

Janzen DH (1967) Why mountain passe are higher in the tropics. Am Nat **101**: 233–249

Ji CJ, Yang YH, Han WX, He YF, Smith J, Smith P (2014) Climatic and Edaphic Controls on Soil pH in Alpine Grasslands on the Tibetan Plateau, China: A Quantitative Analysis. Pedosphere **24**: 39–44
Johnson DB (1961) The spectrophotometric determiantion of lignin in small wood samples. Tappi **44**: 793–798
José Carli (2017) Lignification in *Miscanthus*: biochemical and molecular approaches to furthering understanding. Aberystwyth University
Josse J, Husson F (2016) missMDA: A package for handling missing values in multivariate data analysis. J Stat Softw. doi: 10.18637/jss.v070.i01
Josse J, Husson F (2012) Handling missing values in exploratory multivariate data analysis methods. J la Société Française Stat **153**: 79-99–99
Kahle D, Wickham H (2013) ggmap: Spatial visualization with ggplot2. R J **5**: 144–161
Kalaji HM, B>□ba W, Gediga K, Goltsev V, Samborska IA, Cetner MD, Dimitrova S, Piszcz U, Bielecki K, Karmowska K, et al (2018a) Chlorophyll fluorescence as a tool for nutrient status identification in rapeseed plants. Photosynth Res **136**: 329–343
Kalaji HM, Oukarroum A, Alexandrov V, Kouzmanova M, Brestic M, Zivcak M, Samborska IA, Cetner MD, Allakhverdiev SI, Goltsev V (2014a) Identification of nutrient deficiency in maize and tomato plants by invivo chlorophyll a fluorescence measurements. Plant Physiol Biochem **81**: 16–25
Kalaji HM, Rastogi A, Živčák M, Brestic M, Daszkowska-Golec A, Sitko K, Alsharafa KY, Lotfi R, Stypiński P, Samborska IA, et al (2018b) Prompt chlorophyll fluorescence as a tool for crop phenotyping: an example of barley landraces exposed to various abiotic stress factors. Photosynthetica. doi: 10.1007/s11099-018-0766-z
Kalaji HM, Schansker G, Ladle RJ, Goltsev V, Bosa K, Allakhverdiev SI, Brestic M, Bussotti F, Calatayud A, Dąbrowski P, et al (2014b) Frequently asked questions about in vivo chlorophyll fluorescence: Practical issues. Photosynth Res **122**: 121–158
Kalinina O, Nunn C, Sanderson R, Hastings AFS, Van Der Weijde T, Özgüven M, Tarakanov I, Schüle H, Trindade LM, Dolstra O, et al (2017) Extending *Miscanthus* cultivation with novel germplasm at six contrasting sites. Front Plant Sci **8**: 1–15
Kang S, Post WM, Nichols JA, Wang D, West TO, Bandaru V, Izaurralde RC (2013) Marginal Lands: Concept, Assessment and Management. J Agric Sci **5**: 129–139
Kao W-YY, Tsai T-TT, Chen W-HH (1998) A Comparative Study of *Miscanthus* floridulus (Labill) Warb and M. transmorrisonensis Hayata: Photosynthetic Gas Exchange, Leaf Characteristics and Growth in Controlled Environments. Ann Bot **81**: 295–299
Katō K, Komorita H (1968) Pyrolysis of cellulose. Agric Biol Chem **32**: 21–26
Kawamoto H, Horigoshi S, Saka S (2007) Effects of side-chain hydroxyl groups on pyrolytic β-ether cleavage of phenolic lignin model dimer. J Wood Sci **53**: 268–271
Kelly, K. L. (1965). Twenty two colors of maximum contrast.
Kiemle SN, Zhang X, Esker AR, Toriz G, Gatenholm P, Cosgrove DJ (2014) Role of (1,3)(1,4)-β-glucan in cell walls: Interaction with cellulose. Biomacromolecules **15**: 1727–1736

Kohli K, Prajapati R, Sharma BK (2019) Bio-based chemicals from renewable biomass for integrated biorefineries. Energies. doi: 10.3390/en12020233
Kok MV, Özgür E (2013) Thermal analysis and kinetics of biomass samples. Fuel Process Technol **106**: 739–743
Kruyeniski J, Ferreira PJT, Videira Sousa Carvalho M da G, Vallejos ME, Felissia FE, Area MC (2019) Physical and chemical characteristics of pretreated slash pine sawdust influence its enzymatic hydrolysis. Ind Crops Prod **130**: 528–536
Landi S, Esposito S (2017) Nitrate Uptake Affects Cell Wall Synthesis and Modeling. Front Plant Sci. doi: 10.3389/fpls.2017.01376
Langsrud Ø (2003) ANOVA for unbalanced data: Use Type II instead of Type III sums of squares. Stat Comput **13**: 163–167
Lavoisier 1743-1794 AL (1790) Elements of chemistry : in a new systematic order, containing all the modern discoveries. Edinburgh : W. Creech, 1790.
Lee M-S, Wycislo A, Guo J, Lee DK, Voigt T (2017) Nitrogen Fertilization Effects on Biomass Production and Yield Components of *Miscanthus* ×giganteus. Front Plant Sci **8**: 1–9
Lewandowski I, Clifton-Brown JC, Scurlock JMO, Huisman W (2000) *Miscanthus*: European experience with a novel energy crop. Biomass and Bioenergy **19**: 209–227
Lewandowski I, Scurlock J M O, Lindvall E, Christou M (2003). The development and current status of perennial rhizomatous grasses as energy crops in the US and Europe. Biomass and Bioenergy, 25(4), 335–361. https://doi.org/10.1016/S0961-9534(03)00030-8
Li F, Ren S, Zhang W, Xu Z, Xie G, Chen Y, Tu Y, Li Q, Zhou S, Li Y, et al (2013) Arabinose substitution degree in xylan positively affects lignocellulose enzymatic digestibility after various NaOH/H2SO4 pretreatments in *Miscanthus*. Bioresour Technol **130**: 629–637
Li H, Xue Y, Wu J, Wu H, Qin G, Li C, Ding J, Liu J, Gan L, Long M (2016) Enzymatic hydrolysis of hemicelluloses from *Miscanthus* to monosaccharides or xylo-oligosaccharides by recombinant hemicellulases. Ind Crops Prod **79**: 170–179
Little RJA, Rubin DB, Richards LE, Little RJA, Rubin DB (2002) Statistical Analysis with Missing Data. doi: 10.2307/3172915
Liu C, Xiao L, Jiang J, Wang W, Gu F, Song D, Yi Z, Jin Y, Li L (2013) Biomass properties from different *Miscanthus* species. Food Energy Secur **2**: 12–19
Lu HB, Zhang GY, Jia CX (2009) Analysis on TG-FTIR and kinetics of biomass pyrolysis. 1st Int Conf Sustain Power Gener Supply, SUPERGEN '09 1–5
Lu M, Li J, Han L, Xiao W (2019) An aggregated understanding of cellulase adsorption and hydrolysis for ball-milled cellulose. Bioresour Technol **273**: 1–7
Lygin A V, Upton J, Dohleman FG, Juvik J, Zabotina OA, Widholm JM, Lozovaya V V (2011) Composition of cell wall phenolics and polysaccharides of the potential bioenergy crop –*Miscanthus*. GCB Bioenergy **3**: 333–345
Ma HZ, Cai Z, Zhang FM, Zhang H, Ge S, Dai SL, Chen WL (2015) Taxonomic evaluation of *Miscanthus* nudipes (Poaceae) based on morphological and molecular evidence. Phytotaxa **205**: 1–20
Maathuis FJ (2009) Physiological functions of mineral macronutrients. Curr Opin Plant Biol **12**: 250□t□258
Maathuis FJM (2013) Plant Mineral Nutrients : Methods and Protocols. Humana

Press, Totowa, NJ
Makino Y, Ueno O (2018) Structural and physiological responses of the C4 grass Sorghum bicolor to nitrogen limitation. Plant Prod Sci **21**: 39–50
Mangold A, Lewandowski I, Möhring J, Clifton-Brown J, Krzyżak J, Mos M, Pogrzeba M, Kiesel A (2019) Harvest date and leaf:stem ratio determine methane hectare yield of *Miscanthus* biomass. GCB Bioenergy **11**: 21–33
Marcia M de, Buanafina O, Marcia M. de OB (2009) Feruloylation in grasses: Current and future perspectives. Mol Plant **2**: 861–872
Marini A, Berbenni V, Flor G (1979) Kinetic Parameters from Thermogravimetric Data. 661–663
Marschner P, Rengel Z (2011) Nutrient Availability in Soils. Marschner's Miner. Nutr. High. Plants Third Ed. pp 315–330
Marzec-Schmidt K, Ludwików A, Wojciechowska N, Kasprowicz-Maluśki A, Mucha J, Bagniewska-Zadworna A (2019). Xylem Cell Wall Formation in Pioneer Roots and Stems of Populus trichocarpa (Torr. & Gray). Frontiers in Plant Science, 10, 1419. https://doi.org/10.3389/fpls.2019.01419
Massignam AM, Chapman SC, Hammer GL, Fukai S (2011) Effects of nitrogen supply on canopy development of maize and sunflower. Crop Pasture Sci **62**: 1045–1055
McCalmont JP, Hastings A, McNamara NP, Richter GM, Robson P, Donnison IS, Clifton-Brown J (2015) Environmental costs and benefits of growing *Miscanthus* for bioenergy in the UK. GCB Bioenergy **9**: 489–507
McCann MC, Carpita NC (2015) Biomass recalcitrance: a multi-scale, multi-factor, and conversion-specific property. J Exp Bot **66**: 4109–4118
McCann MC, Wells B, Roberts K (1990) Direct visualization of cross-links in the primary plant cell wall. J Cell Sci **96**: 323–334
Mehmood MA, Ibrahim M, Rashid U, Nawaz M, Ali S, Hussain A, Gull M (2017) Biomass production for bioenergy using marginal lands. Sustain Prod Consum **9**: 3–21
Mettler MS, Paulsen AD, Vlachos DG, Dauenhauer PJ (2012) The chain length effect in pyrolysis: Bridging the gap between glucose and cellulose. Green Chem **14**: 1284–1288
Monti A, Cosentino SL (2015) Conclusive Results of the European Project OPTIMA: Optimization of Perennial Grasses for Biomass Production in the Mediterranean Area. BioEnergy Res **8**: 1459–1460
Müller J, Toev T, Heisters M, Teller J, Moore KL, Hause G, Dinesh DC, Bürstenbinder K, Abel S (2015) Iron-Dependent Callose Deposition Adjusts Root Meristem Maintenance to Phosphate Availability. Dev Cell **33**: 216–230
Neumann G, Römheld V (2011) Rhizosphere Chemistry in Relation to Plant Nutrition. Marschner's Miner Nutr High Plants Third Ed 347–368
Le Ngoc Huyen T, Remond C, Dheilly RM, Chabbert B (2010) Effect of harvesting date on the composition and saccharification of *Miscanthus* x giganteus. Bioresour Technol **101**: 8224–8231
Nishiyama Y, Kim UJ, Kim DY, Katsumata KS, May RP, Langan P (2003) Periodic disorder along ramie cellulose microfibrils. Biomacromolecules **4**: 1013–1017
Nothnagel EA, Mcneil M, Albersheim P, Dell A (1983) Host-Pathogen Interactions. 916–926
Ogden M, Hoefgen R, Roessner U, Persson S, Khan GA (2018) Feeding the

walls: How does nutrient availability regulate cellwall composition? Int J Mol Sci. doi: 10.3390/ijms19092691

Okunlola GO, Adelusi AA (2014). Growth and Photosynthetic Pigment Accumulation in Lycopersicum esculentum Mill. In Response to Light and Nutrient Stress. Notulae Scientia Biologicae, 6(2), 250–255. https://doi.org/10.15835/nsb.6.2.9195

Olsen KM, Purugganan MD (2002) Molecular evidence on the origin and evolution of glutinous rice. Genetics **162**: 941–950

Palaisa K, Morgante M, Tingey S, Rafalski A (2004) Long-range patterns of diversity and linkage disequilibrium surrounding the maize Y1 gene are indicative of an asymmetric selective sweep. Proc Natl Acad Sci U S A **101**: 9885–9890

Pancaldi F, Trindade LM (2020) Marginal Lands to Grow Novel Bio-Based Crops : A Plant Breeding Perspective. **11**: 1–24

Pang CH, Gaddipatti S, Tucker G, Lester E, Wu T (2014) Relationship between thermal behaviour of lignocellulosic components and properties of biomass. Bioresour Technol **172**: 312–320

Pattathil S, Ingwers MW, Victoriano OL, Kandemkavil S, McGuire MA, Teskey RO, Aubrey DP (2016) Cell wall ultrastructure of stem wood, roots, and needles of a conifer varies in response to moisture availability. Front Plant Sci **7**: 1–11

Peirson BRE (2015) Plasticity, stability, and yield: The origins of Anthony David Bradshaw's model of adaptive phenotypic plasticity. Stud Hist Philos Sci Part C Stud Hist Philos Biol Biomed Sci **50**: 51–66

Peña MJ, Kulkarni AR, Backe J, Boyd M, O'Neill MA, York WS (2016) Structural diversity of xylans in the cell walls of monocots. Planta **244**: 589–606

Perrier L, Rouan L, Jaffuel S, Clément-Vidal A, Roques S, Soutiras A, Baptiste C, Bastianelli D, Fabre D, Dubois C, et al (2017) Plasticity of sorghum stem biomass accumulation in response to water deficit: A multiscale analysis from internode tissue to plant level. Front Plant Sci **8**: 1–14

Peterson GM, Galbraith JK (1932) The Concept of Marginal Land. J Farm Econ **14**: 295

Petit J, Gulisano A, Dechesne A, Trindade LM (2019) Phenotypic Variation of Cell Wall Composition and Stem Morphology in Hemp (Cannabis sativa L.): Optimization of Methods. Front Plant Sci. doi: 10.3389/fpls.2019.00959

Pettolino FA, Walsh C, Fincher GB, Bacic A (2012) Determining the polysaccharide composition of plant cell walls. Nat Protoc **7**: 1590–1606

Piperno DR, Moreno JE, Iriarte J, Holst I, Lachniet M, Jones JG, Ranere AJ, Castanzo R (2007) Late Pleistocene and Holocene environmental history of the Iguala Valley, Central Balsas Watershed of Mexico. Proc Natl Acad Sci **104**: 11874 LP – 11881

van der Ploeg RR, Böhm W, Kirkham MB (1999) On the Origin of the Theory of Mineral Nutrition of Plants and the Law of the Minimum. Soil Sci Soc Am J **63**: 1055–1062

Podgórska A, Burian M, Gieczewska K, Ostaszewska-Bugajska M, Zebrowski J, Solecka D, Szal B (2017). Altered Cell Wall Plasticity Can Restrict Plant Growth under Ammonium Nutrition. Frontiers in Plant Science, 8, 1344. https://doi.org/10.3389/fpls.2017.01344

Prasad V (2005) Dinosaur Coprolites and the Early Evolution of Grasses and Grazers. Science (80-) **310**: 1177–1180

Purugganan MD, Fuller DQ (2009) The nature of selection during plant domestication. Nature **457**: 843–848
Quinn LD, Straker KC, Guo J, Kim S, Thapa S, Kling G, Lee DK, Voigt TB (2015) Stress-Tolerant Feedstocks for Sustainable Bioenergy Production on Marginal Land. Bioenergy Res **8**: 1081–1100
Quinto-Mosquera H, Moreno-Hurtado F (2016) Precipitation effects on soil characteristics in tropical rain forests of the chocó biogeographical region. Rev Fac Nac Agron Medellin **69**: 7813–7823
R Core Team (2018) R: A Language and Environment for Statistical Computing.
Ralph J, Quideau S, Grabber JH, Hatfield RD (1994) Identification and synthesis of new ferulic acid dehydrodimers present in grass cell walls. J Chem Soc Perkin Trans 1 **602**: 3485
Ramírez V, Xiong G, Mashiguchi K, Yamaguchi S, Pauly M (2018) Growth- and stress-related defects associated with wall hypoacetylation are strigolactone-dependent. Plant Direct. doi: 10.1002/pld3.62
Rancour DM, Marita JM, Hatfield RD (2012) Cell wall composition throughout development for the model grass Brachypodium distachyon. Front Plant Sci **3**: 1–14
Ratnayake S, Beahan CT, Callahan DL, Bacic A (2014) The reducing end sequence of wheat endosperm cell wall arabinoxylans. Carbohydr Res **386**: 23–32
Reed HS (1942) A short history of the plant sciences. 5 p. l., 320, [3]
Rennenberg H, Schmidt S (2010). Perennial lifestyle—An adaptation to nutrient limitation? Tree Physiology, 30(9), 1047–1049. https://doi.org/10.1093/treephys/tpq076
Resch MG, Baker JO, Decker SR (2015) Low Solids Enzymatic Saccharification of Lignocellulosic Biomass. Tech Rep NREL/TP-5100-63351, Lab Anal Proced 1–9
Robertson GP, Hamilton SK, Barham BL, Dale BE, Izaurralde RC, Jackson RD, Landis DA, Swinton SM, Thelen KD, Tiedje JM (2017) Cellulosic biofuel contributions to a sustainable energy future: Choices and outcomes. Science (80-). doi: 10.1126/science.aal2324
Robinson AR, Mansfield SD (2009) Rapid analysis of poplar lignin monomer composition by a streamlined thioacidolysis procedure and near-infrared reflectance-based prediction modeling. Plant J **58**: 706–714
Robson P, Jensen E, Hawkins S, White SR, Kenobi K, Clifton-Brown J, Donnison I, Farrar K (2013) Accelerating the domestication of a bioenergy crop: Identifying and modelling morphological targets for sustainable yield increase in *Miscanthus*. J Exp Bot **64**: 4143–4155
Rodríguez D, Keltjens WG, Goudriaan J (1998) Plant leaf area expansion and assimilate production in wheat (Triticum aestivum L.) growing under low phosphorus conditions. Plant Soil **200**: 227–240
Saeman JF (1945) Kinetics of Wood Saccharification - Hydrolysis of Cellulose and Decomposition of Sugars in Dilute Acid at High Temperature. Ind Eng Chem **37**: 43–52
Saldarriaga JF, Aguado R, Pablos A, Amutio M, Olazar M, Bilbao J (2015) Fast characterization of biomass fuels by thermogravimetric analysis (TGA). Fuel **140**: 744–751
Santoro N, Cantu SL, Tornqvist CE, Falbel TG, Bolivar JL, Patterson SE, Pauly M, Walton JD (2010) A high-throughput platform for screening

milligram quantities of plant biomass for lignocellulose digestibility. Bioenergy Res **3**: 93–102

Santos RB, Lee JM, Jameel H, Chang HM, Lucia LA (2012) Effects of hardwood structural and chemical characteristics on enzymatic hydrolysis for biofuel production. Bioresour Technol **110**: 232–238

Santos VT de O, Siqueira G, Milagres AMF, Ferraz A (2018) Role of hemicellulose removal during dilute acid pretreatment on the cellulose accessibility and enzymatic hydrolysis of compositionally diverse sugarcane hybrids. Ind Crops Prod **111**: 722–730

Schäfer J, Sattler M, Iqbal Y, Lewandowski I, Bunzel M (2019) Characterization of *Miscanthus* cell wall polymers. GCB Bioenergy **11**: 191–205

Schertz F. M. (1929) The effect of potassium, nitrogen and phosphorus fertilizing upon the chloroplast pigments, upon the mineral content of the leaves, and upon production in crop plants. Plant Physiol **4**: 269–279

Sederoff RR, MacKay JJ, Ralph J, Hatfield RD (1999) Unexpected variation in lignin. Curr Opin Plant Biol **2**: 145–152

Sheng F, Xiuling C (2004) Influence of Atmospheric Precipitation on Soil Leaching and Desalinization in Semi-Humid Region. TucsonArsAgGov 1–4

Sher F, Iqbal SZ, Liu H, Imran M, Snape CE (2019) Thermal and kinetic analysis of diverse biomass fuels under different reaction environment: A way forward to renewable energy sources. Energy Convers Manag **203**: 112266

Singh B, Schulze DG (2015) Soil Minerals and Plant Nutrition. Nat Educ Knowl **6**: 1

Singh SK, Reddy VR, Fleisher DH, Timlin DJ (2019) Interactive effects of temperature and phosphorus nutrition on soybean: Leaf photosynthesis, chlorophyll fluorescence, and nutrient efficiency. Photosynthetica **57**: 248–257

Slavov G, Robson P, Jensen E, Hodgson E, Farrar K, Allison G, Hawkins S, Thomas-Jones S, Ma X-F, Huang L, et al (2013) Contrasting geographic patterns of genetic variation for molecular markers vs. phenotypic traits in the energy grass*Miscanthus* sinensis. GCB Bioenergy **5**: 562–571

Slessarev EW, Lin Y, Bingham NL, Johnson JE, Dai Y, Schimel JP, Chadwick OA (2016) Water balance creates a threshold in soil pH at the global scale. Nature **540**: 567–569

Sluiter a, Hames B, Hyman D, Payne C, Ruiz R, Scarlata C, Sluiter J, Templeton D, Nrel JW (2008) Determination of total solids in biomass and total dissolved solids in liquid process samples. Natl Renew Energy Lab 9

Sluiter A, Hames B, Ruiz R, Scarlata C, Sluiter J, Templeton D, Crocker D (2012) NREL/TP-510-42618 analytical procedure - Determination of structural carbohydrates and lignin in Biomass. Lab Anal Proced 17

Smart AJ, Schacht WH, Pedersen JF, Undersander DJ, Moser LE (1998) Prediction reflectance of leaf : stem ratio in grasses spectroscopy using near infrared. **51**: 447–449

Sprenger C (1839) Die Lehre vom Dünger oder Beschreibung aller bei der Landwirthschaft gebräuchlicher vegetablilischer, animalischer und mineralischer Düngermaterialien, nebst Erklärung ihrer Wirkungsart. Liepzig

Stavridou E, Hastings A, Webster RJ, Robson PRH (2016) The impact of soil salinity on the yield, composition and physiology of the bioenergy grass *Miscanthus* x giganteus. GCB Bioenergy n/a-n/a

Stavridou E, Webster RJ, Robson PRH (2019) Novel *Miscanthus* genotypes

selected for different drought tolerance phenotypes show enhanced tolerance across combinations of salinity and drought treatments. Ann Bot **124**: 653–674

Stocker DF, Qin D, Plattner G-K, Tignor M, S.K. A, J. B, Nauels A, Xia Y, Bex V, Midgley PM (2013) Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, 2013 Summ Policymakers. doi: 10.1260/095830507781076194

Strasser RJ, Srivastava A, Tsimilli-Michael M (2000) The fluorescence transient as a tool to characterize and screen photosynthetic samples. Probing Photosynth Mech Regul Adapt 443–480

Sun Q, Lin Q, Yi Z-L, Yang Z-R, Zhou F-S (2010) A taxonomic revision of *Miscanthus* s.l. (Poaceae) from China. Bot J Linn Soc **164**: 178–220

Sunday JM, Bates AE, Kearney MR, Colwell RK, Dulvy NK, Longino JT, Huey RB (2014) Thermal-safety margins and the necessity of thermoregulatory behavior across latitude and elevation. Proc Natl Acad Sci U S A **111**: 5610–5615

Suntivarakorn R, Treedet W, Singbua P, Teeramaetawat N (2018) Fast pyrolysis from Napier grass for pyrolysis oil production by using circulating Fluidized Bed Reactor: Improvement of pyrolysis system and production cost. Energy Reports **4**: 565–575

Takkellapati S, Li T, Gonzalez MA (2018) An overview of biorefinery-derived platform chemicals from a cellulose and hemicellulose biorefinery. Clean Technol Environ Policy **20**: 1615–1630

Tam RK, Magistad OC (1935) Relationship between Nitrogen Fertilization and Chlorophyll Content in Pineapple Plants. Plant Physiol **10**: 159–168

Tamura K ichi, Uwatoko N, Yamashita H, Fujimori M, Akiyama Y, Shoji A, Sanada Y, Okumura K, Gau M (2016) Discovery of Natural Interspecific Hybrids Between *Miscanthus* Sacchariflorus and *Miscanthus* Sinensis in Southern Japan: Morphological Characterization, Genetic Structure, and Origin. Bioenergy Res **9**: 315–325

Tang C, Yang X, Chen X, Ameen A, Xie G (2018) Sorghum biomass and quality and soil nitrogen balance response to nitrogen rate on semiarid marginal land. F Crop Res **215**: 12–22

Tang Y, Horikoshi M, Li W (2016) Ggfortify: Unified interface to visualize statistical results of popular r packages. R J **8**: 478–489

Tejera MD, Heaton EA (2017). Description and Codification of *Miscanthus* × giganteus Growth Stages for Phenological Assessment. Frontiers in Plant Science, 8(October), 1–12. https://doi.org/10.3389/fpls.2017.01726

Tenhaken R (2014) Cell wall remodeling under abiotic stress. Front Plant Sci **5**: 771

The Plant List (2013) The Plant List. Version 1.1, http://www.theplantlist.org/

Tian L, Shen B, Xu H, Li F, Wang Y, Singh S (2016) Thermal behavior of waste tea pyrolysis by TG-FTIR analysis. Energy **103**: 533–542

Tonzig S, Marré E (1983) Botanica Generale - Morfologia e fisiologia vegetali. Casa Editrice Ambrosiana, Milano

Tufte 1942- ER The visual display of quantitative information. Second edition. Cheshire, Conn. : Graphics Press, [2001] ©2001

Tufte ER (2001) The Visual Display of Quantitative Information Aesthetics and Technique in Data Graphical Design, Second. Graphic Press LLC, Cheshire, Connecticut

Tukey JW (1949) Comparing Individual Means in the Analysis of Variance Author (

s): John W . Tukey Published by : International Biometric Society Stable URL : http://www.jstor.org/stable/3001913. Int Biometric Soc **5**: 99–114
Updegraff DM (1969) Semimicro determination of cellulose in biological materials. Anal Biochem **32**: 420–424
Usuda H (1995) Phosphate deficiency in maize. V. Mobilization of nitrogen and phosphorus within shoots of young plants and its relationship to senescence. Plant Cell Physiol **36**: 1041–1049
Uwatoko N, Tamura K, Yamashita H, Gau M (2016) Naturally occurring triploid hybrids between *Miscanthus* sacchariflorus and *M. sinensis* in Southern Japan, show phenotypic variation in agronomic and morphological traits. Euphytica 1–16
van der Weijde T, Alvim Kamei CL, Torres AF, Vermerris W, Dolstra O, Visser RG, Trindade LM (2013) The potential of C4 grasses for cellulosic biofuel production. Front Plant Sci **4**: 107
van der Weijde T, Dolstra O, Visser RGF, Trindade LM (2017a) Stability of cell wall composition and saccharification efficiency in *Miscanthus* across diverse environments. Front Plant Sci **7**: 1–14
van der Weijde T, Kiesel A, Iqbal Y, Muylle H, Dolstra O, Visser RGFF, Lewandowski I, Trindade LM (2017b) Evaluation of *Miscanthus* sinensis biomass quality as feedstock for conversion into different bioenergy products. GCB Bioenergy **9**: 176–190

van der Weijde T, Huxley LM, Hawkins S, Sembiring EH, Farrar K, Dolstra O, Visser RGF, Trindade LM (2016) Impact of drought stress on growth and quality of *Miscanthus* for biofuel production. GCB Bioenergy n/a-n/a

Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W (2010). Lignin Biosynthesis and Structure. Plant Physiology, 153(3), 895–905. https://doi.org/10.1104/pp.110.155119
Van Soest PJ (1963) Use of detergents in the analysis of fibrous feeds. 2. A rapid method for the determination of fiber and lignin. J Assoc Off Agric Chem **46**: 829–835
Various (1882) The Linnean Society. J Linn Soc Bot. doi: 10.1038/144664e0
Varnero CS, Urrutia MV, Ibaceta SV (2018) Bioenergy from Perennial Grasses. Adv. Biofuels Bioenergy. InTech, p 13
Vazquez De Aldana BR, Berendse F (1997). Nitrogen-use efficiency in six perennial grasses from contrasting habitats. Functional Ecology, 11(5), 619–626. https://doi.org/10.1046/j.1365-2435.1997.00137.x
Venderbosch RH, Prins W (2011) Fast Pyrolysis. Thermochem. Process. Biomass Convers. into Fuels, Chem. Power. pp 124–156
Verbanck M, Josse J, Husson F (2013) Regularised PCA to denoise and visualise data. Stat Comput **25**: 471–486
Voesenek LACJ, Van Veen H, Sasidharan R (2014) Learning from nature: The use of non-model species to identify novel acclimations to flooding stress. AoB Plants **6**: 1–6
Voragen AGJ, Coenen G-J, Verhoef RP, Schols HA (2009) Pectin, a versatile polysaccharide present in plant cell walls. Struct Chem **20**: 263
Walker WR (1989) Guidelines for designing and evaluating surface irrigation systems.

Wang S, Dai G, Yang H, Luo Z (2017) Lignocellulosic biomass pyrolysis mechanism: A state-of-the-art review. Prog Energy Combust Sci **62**: 33–86
Wang WL, Ren XY, Che YZ, Chang JM, Gou JS (2013) Kinetics and FTIR characteristics of the pyrolysis process of poplar wood. For Sci Pract **15**: 70–75
Wang X, Tang Q, Zhao X, Jia C, Yang X, He G, Wu A, Kong Y, Hu R, Zhou G (2016a) Functional conservation and divergence of *Miscanthus* lutarioriparius GT43 gene family in xylan biosynthesis. BMC Plant Biol **16**: 102
Wang Y, Fan C, Hu H, Li Y, Sun D, Wang Y, Peng L (2016b) Genetic modification of plant cell walls to enhance biomass yield and biofuel production in bioenergy crops. Biotechnol Adv. doi: http://dx.doi.org/10.1016/j.biotechadv.2016.06.001
Wang Z, McDonald AG, Westerhof RJM, Kersten SRA, Cuba-Torres CM, Ha S, Pecha B, Garcia-Perez M (2013b) Effect of cellulose crystallinity on the formation of a liquid intermediate and on product distribution during pyrolysis. J Anal Appl Pyrolysis **100**: 56–66
Weih M, Asplund L, Bergkvist G (2011) Assessment of nutrient use in annual and perennial crops: A functional concept for analyzing nitrogen use efficiency. Plant Soil **339**: 513–520
Weih M, Hamnér K, Pourazari F (2018) Analyzing plant nutrient uptake and utilization efficiencies: comparison between crops and approaches. Plant Soil **430**: 7□:21
Werpy T, Petersen G (2004) Top Value Added Chemicals from Biomass: Volume I -- Results of Screening for Potential Candidates from Sugars and Synthesis Gas. doi: 10.2172/15008859
White PJ (2011) Ion Uptake Mechanisms of Individual Cells and Roots: Short-distance Transport. Marschner's Miner Nutr High Plants Third Ed. doi: 10.1016/B978-0-12-384905-2.00002-9
Wickham H (2011) The split-apply-combine strategy for data analysis. J Stat Softw **40**: 1–29
Wickham H (2016) ggplot2: Elegant graphics for data analysis. Springer
Wicklam H, Romain F, Henry L, Müller K (2019) dplyr: A Grammar of Data Manipulation. R package version 0.8.3. CRAN, https://cran.r-project.org/package=dplyr
Wieczorek J, Bloom D, Guralnick R, Blum S, Dr□ring M, Giovanni R, Robertson T, Vieglais D (2012) Darwin core: An evolving community-developed biodiversity data standard. PLoS One. doi: 10.1371/journal.pone.0029715
Wu HC, Bulgakov VP, Jinn TL (2018) Pectin methylesterases: Cell wall remodeling proteins are required for plant response to heat stress. Front Plant Sci **871**: 1–21
Wu Y wei, Li Q, Jin R, Chen W, Liu X lin, Kong F lei, Ke Y pei, Shi H chun, Yuan J chao (2019) Effect of low-nitrogen stress on photosynthesis and chlorophyll fluorescence characteristics of maize cultivars with different low-nitrogen tolerances. J Integr Agric **18**: 1246–1256
Wu Z, Zhang M, Wang L, Tu Y, Zhang J, Xie G, Zou W, Li F, Guo K, Li Q, et al (2013) Biomass digestibility is predominantly affected by three factors of wall polymer features distinctive in wheat accessions and rice mutants. Biotechnol Biofuels **6**: 183
Xin D, Chen X, Wen P, Zhang J (2019). Insight into the role of α-arabinofuranosidase

in biomass hydrolysis: Cellulose digestibility and inhibition by xylooligomers. Biotechnology for Biofuels, 12(1), 1–11. https://doi.org/10.1186/s13068-019-1412-0

Xiong Q, Tang G, Zhong L, He H, Chen X (2018) Response to nitrogen deficiency and compensation on physiological characteristics, yield formation, and nitrogen utilization of rice. Front Plant Sci **9**: 1–14

Xu F xiang, Zhang X, Zhang F, Jiang L qun, Zhao Z li, Li H bin (2020) TG-FTIR for kinetic evaluation and evolved gas analysis of cellulose with different structures. Fuel **268**: 117365

Xu N, Zhang W, Ren S, Liu F, Zhao C, Liao H, Xu Z, Huang J, Li Q, Tu Y, Yu B, Wang Y, Jiang J, Qin J, Peng L (2012). Hemicelluloses negatively affect lignocellulose crystallinity for high biomass digestibility under NaOH and H2SO4 pretreatments in *Miscanthus*. Biotechnology for Biofuels, 5(1), 1–12. https://doi.org/10.1186/1754-6834-5-58

Xu HX, Weng XY, Yang Y (2007) Effect of phosphorus deficiency on the photosynthetic characteristics of rice plants. Russ J Plant Physiol **54**: 741–748

Xue J, Bosch M, Knox JP (2013) Heterogeneity and glycan masking of cell wall microstructures in the stems of *Miscanthus* x giganteus, and its parents *M. sinensis* and *M. sacchariflorus*. PLoS One **8**: e82114

Yan B, Ji Z, Fan B, Wang X, He G, Shi L, Liu G (2016) Plants adapted to nutrient limitation allocate less biomass into stems in an arid-hot grassland. New Phytol **211**: 1232–1240

Yan J, Chen W, Luo F, Ma H, Meng A, Li X, Zhu M, Li S, Zhou H, Zhu W, et al (2012) Variability and adaptability of *Miscanthus* species evaluated for energy crop domestication. GCB Bioenergy **4**: 49–60

Yan J, Zhu C, Liu W, Luo F, Mi J, Ren Y, Li J, Sang T (2015) High photosynthetic rate and water use efficiency of *Miscanthus* lutarioriparius characterize an energy crop in the semiarid temperate region. GCB Bioenergy **7**: 207–218

Yang H, Yan R, Chen H, Lee DH, Zheng C (2007) Characteristics of hemicellulose, cellulose and lignin pyrolysis. Fuel **86**: 1781–1788

Yang H, Zhang X, Luo H, Liu B, Shiga TM, Li X, Kim JI, Rubinelli P, Overton JC, Subramanyam V, et al (2019) Overcoming cellulose recalcitrance in woody biomass for the lignin-first biorefinery. Biotechnol Biofuels **12**: 171

Yoo CG, Yang Y, Pu Y, Meng X, Muchero W, Yee KL, Thompson OA, Rodriguez M, Bali G, Engle NL, et al (2017a) Insights of biomass recalcitrance in natural Populus trichocarpa variants for biomass conversion. Green Chem **19**: 5467–5478

Yoo J-S, Kim C-H, Yoon Y-M (2017b) Biochemical Methane Potential Analysis for Anaerobic Digestion of Giant *Miscanthus* (*Miscanthus* sacchariflorus). Korean J Environ Agric **36**: 29–35

van der Zee SEATM, Shah SHH, Vervoort RW (2014) Root zone salinity and sodicity under seasonal rainfall due to feedback of decreasing hydraulic conductivity. Water Resour Res **50**: 9432–9446

Zhang H, Fangel JU, Willats WGTT, Selig MJ, Lindedam J, Jørgensen H, Felby C (2014a) Assessment of leaf/stem ratio in wheat straw feedstock and impact on enzymatic conversion. GCB Bioenergy **6**: 90–96

Zhang J, Feng L, Wang D, Zhang R, Liu G, Cheng G (2014b) Thermogravimetric

analysis of lignocellulosic biomass with ionic liquid pretreatment. Bioresour Technol **153**: 379–382

Zhao H, Li Q, He J, Yu J, Yang J, Liu C, Peng, J (2014). Genotypic variation of cell wall composition and its conversion efficiency in *Miscanthus* sinensis , a potential biomass feedstock crop in China. GCB Bioenergy, 6(6), 768–776. https://doi.org/10.1111/gcbb.12115

Zivcak M, Brestic M, Kunderlikova K, Olsovska K, Allakhverdiev SI (2015) Effect of photosystem i inactivation on chlorophyll a fluorescence induction in wheat leaves: Does activity of photosystem i play any role in OJIP rise? J Photochem Photobiol B Biol **152**: 318–324

Živčák M, Olšovská K, Slamka P, Galambošová J, Rataj V, Shao HB, Brestič M (2014a) Application of chlorophyll fluorescence performance indices to assess the wheat photosynthetic functions influenced by nitrogen deficiency. Plant, Soil Environ **60**: 210–215

Živčák M, Olšovská K, Slamka P, Galambošová J, Rataj V, Shao HB, Kalaji HM, Brestič M (2014b) Measurements of chlorophyll fluorescence in different leaf positions may detect nitrogen deficiency in wheat. Zemdirbyste **101**: 437–444

Zoghlami A, Paës G (2019). Lignocellulosic Biomass: Understanding Recalcitrance and Predicting Hydrolysis. Frontiers in Chemistry, 7, 874. https://doi.org/10.3389/fchem.2019.00874

Appendixes

Appendix A Protocols

A1 Protocols

A1.1 Potassium acetate buffer, 0.1 M

Solution A: 11.55 ml glacial acetic acid per liter (0.2 M) in water.

Solution B: 19.6 g potassium acetate ($KC_2H_3O_2$) per liter (0.2 M) in water

Referring to Table A. 1 for desired pH, mix the indicated volumes of solutions A and B, then dilute with water to 100 ml. Filter sterilize if necessary.

Store up to 3months at room temperature. This may be made as a 5- or 10-fold concentrate by scaling up the amount of sodium acetatein the same volume. Acetate buffers show concentration-dependent pH changes, so check the pH by diluting an aliquot of concentrate to the final concentration. To prepare buffers with pH intermediate between the points listed in Table A.2, prepare closest higher pH, then titrate with solution A.

Table A. 1

Desired pH	Solution A(ml)	Solution B(ml)
3.6	46.3	3.7
3.8	44.0	6.0
4.0	41.0	9.0
4.2	36.8	13.2

4.4	30.5	19.5
4.6	25.5	24.5
4.8	20.0	30.0
5.0	14.8	35.2
5.2	10.5	39.5
5.4	8.8	41.2
5.6	4.8	45.2

Appendix B Chapter 1

B1 Tables

Table A 1 List of the species confirmed species of the genus *Miscanthus*

Family	Genus	hyb mark	Species	Authorship	Publication	Collation	Page	Date
Poaceae	*Miscanthus*		changii	Y.N.Lee	J. Jap. Bot.	39: 115		1964
Poaceae	*Miscanthus*		depauperatus	Merr.	Philipp. J. Sci., C	0.326389		1910
Poaceae	*Miscanthus*		ecklonii	(Nees) Mabb.	Taxon	33: 442		1984
Poaceae	*Miscanthus*		erectus	Gibbs.-Russ.				
Poaceae	*Miscanthus*		floridulus	(Labill.) Warb. ex K.Schum. & Lauterb.	Fl. Schutzgeb. Südsee	166		1900
Poaceae	*Miscanthus*		fuscus	(Roxb.) Benth.	J. Linn. Soc., Bot.	0.836806		1881
Poaceae	*Miscanthus*	×	giganteus	J.M.Greef , Deuter ex Hodk., Renvoize	Kew Bull.	56: 759		2001

Poaceae	*Miscanthus*	junceus	(Stapf) Pilg.	Nat. Pflanzenfam. ed. 2	14e: 113		1940
Poaceae	*Miscanthus*	nepalensis	(Trin.) Hack.	Monogr. Phan.	0.322222		1889
Poaceae	*Miscanthus*	nudipes	(Griseb.) Hack.	Monogr. Phan.	0.325694		1889
Poaceae	*Miscanthus*	oligostachyus	Stapf	Bull. Misc. Inform. Kew	1898: 227		1898
Poaceae	*Miscanthus*	paniculatus	(B.S. Sun) Renvoize & S.L. Chen	Kew Bull.	60(4): 607	607	2005
Poaceae	*Miscanthus*	sacchariflorus	(Maxim.) Hack.	Nat. Pflanzenfam.	2(2): 23		1887
Poaceae	*Miscanthus*	sinensis	Andersson	Öfvers. Förh. Kongl. Svenska Vetensk.-Akad.	0.615278		1855
Poaceae	*Miscanthus*	tinctorius	(Steud.) Hack.	Monogr. Phan.	0.321528		1889
Poaceae	*Miscanthus*	violaceus	(K.Schum.) Pilg.	Nat. Pflanzenfam. ed. 2	14e: 113		1940

Appendix C Chapter 2

C1 Tables

Table C. 1 Genotypes selected from the population in ABR33 field.

The group includes some hybrids artificially created at IBERS. For these latter ones coordinates are not available.

Genotype	N	accession	Name	Genetic group	pop5	Lat	Long	Alt	Country
Mb1181#25	1	Mb1181	Llyr	Flor, Taiwan	4	23.2899	120.498	2282	Taiwan
Mb1029#24	2	Mb1029	Nudd	Sacc rob	3	24.18619	121.3348	2340	Taiwan
Mb1030#216	3	Mb1030	Lancelot	Flor, Taiwan	4	24.06064	121.1552	1932	Taiwan
Mb337	4	Mb337	Modred	Sacc rob	3	34.3045	117.951	12	China
Mb340	5	Mb340	Lunete	Sacc rob	3	34.3057	117.951	12	China
Mb357	6	Mb357	Beli	Sacc rob	3	34.3013	117.945	18	China
Mb348	7	Mb348	Pryderi	Sacc rob	3	34.3048	117.951	24	China
Mb1123#27	8	Mb1123	Gawain	Flor, Taiwan	4	24.6254	121.485	1213	Taiwan
Mb1102#46	9	Mb1102	Geraint	Flor, Taiwan	4	24.0611	121.116	2309	Taiwan

Genotype	N	accession	Name	Genetic group	pop5	Lat	Long	Alt	Country
Mb1116#31	10	Mb1116	Percival	Flor, Taiwan	4	23.4875	120.89	2623	Taiwan
Mb229	11	Mb229	Angharad	Sin South Japan	2	40.91	124.852	450	China
Mb494#1	12	Mb494	Pwyll	Sacc rob	3	33.9408	133.635	222	Japan
Mb25	13	Mb25	Enid	Sin N Japan	1	36.141	137.725	1320	Japan
Mb13	14	Mb13	Peredur	Sin N Japan	1	37.0161	140.121	1260	Japan
Mb215	15	NA	Dylan	Sin N Japan	1	NA	NA	NA	NA
Mb112	16	NA	Luned	Sin (EMI/PRI)	1	NA	NA	NA	NA
Mb102	17	NA	Uther	Sin, Taiwan	2	NA	NA	NA	NA
Mb255	18	NA	Bors	Sin, Taiwan	1	NA	NA	NA	NA
Mb915#19	19	Mb915	Aeron	Sin N Japan	1	35.5094	137.543	460	Japan
Mb1086#54	20	Mb1086	Lamorak	Hyb	5	29.0442	111.415	36	China
Mb1069#920	21	Mb1069	Lleu	Lut	3	28.5335	112.077	33	China
Mb1060#1	22	Mb1060	Eigyr	Lut	3	28.5124	112.248	19	China

Genotype	N	accession	Name	Genetic group	pop5	Lat	Long	Alt	Country
Mb1062#15	23	Mb1062	Blodeuwedd	Lut	3	28.5408	112.224	30	China
Mb1074#21	24	Mb1074	Bran	Lut	3	28.5435	112.075	23	China
Mb1077#44	25	Mb1077	Gwydion	Lut	3	28.5522	112.092	31	China
Mb383#1	26	Mb383	Math	Lut	3	28.8305	112.396	33	China
Mb415#1	27	Mb415	Govannon	Lut	3	28.8824	112.439	30	China
Mb424#8	28	Mb424	Goronwy	Lut	3	29.4283	113.066	22	China
Mb1082#41	29	Mb1082	Gwalchmei	Lut	3	28.5942	112.015	22	China
Mb779#14	30	Mb779	Kay	Lut	3	31.1692	118.221	2	China
Mb627#232	31	Mb627	Mabon	Lut	3	28.88572	112.1019	26	China
Mb426#1	32	Mb426	Culhwch	Sacc rob	3	37.4114	118.657	-1	China
Mb427#1	33	Mb427	Aranrhod	Sacc rob	3	37.4114	118.657	-1	China

Genotype	N	accession	Name	Genetic group	pop5	Lat	Long	Alt	Country
Mb429#1	34	Mb429	Gwenhwyfar	Sacc rob	3	37.4114	118.657	-1	China
Mb1094#900	35	Mb1094	Owain	Lut	3	29.0319	111.345	40	China
Mb358#2	36	Mb358	Rhiannon	Sacc rob	3	34.2988	117.952	27	China
Mb1201#13	37	Mb1201	Urien	Sin S Japan	2	35.4304	133.284	2	Japan
Mb1185#11	38	Mb1185	Gareth	Sin N Japan	1	35.6165	135.974	20	Japan
Mb1205#12	39	Mb1205	Bedivere	Sin S Japan	2	35.2884	132.686	62	Japan
Mb52	40	Mb52	Olwen	Hyb	5	37.9667	128.467	510	Korea, South
Mb87	41	Mb87	Gaheris	Sacc rob	3	35.7333	127.583	350	Korea, South
Mb1328#12	42	Mb1328	Yspaddaden	Sin Taiwan	1	24.1205	121.269	2997	Taiwan
Mb1564#1	43	Mb1564	Eufydd	Flor Taiwan	4	25.14324	121.5206		Taiwan

Genotype	N	accession	Name	Genetic group	pop5	Lat	Long	Alt	Country
Mb311	44	Mb311	Cilydd	Hyb.gig	5	NA	NA	NA	NA
Mx1907#12	45	Mx1907	Cigfa	Hyb.mx	1	NA	NA	NA	NA
Mx1901#1	46	Mx1901	Hafgan	Hyb.mx	1	NA	NA	NA	NA
Mx1968#15	47	Mx1968	Nimue	Hyb.mx	1	NA	NA	NA	NA
Mx2765#5	48	Mx2765	Galahd	Hyb.mx	5	NA	NA	NA	NA
Mx2765#11	49	Mx2765	Tristan	Hyb.mx	5	NA	NA	NA	NA

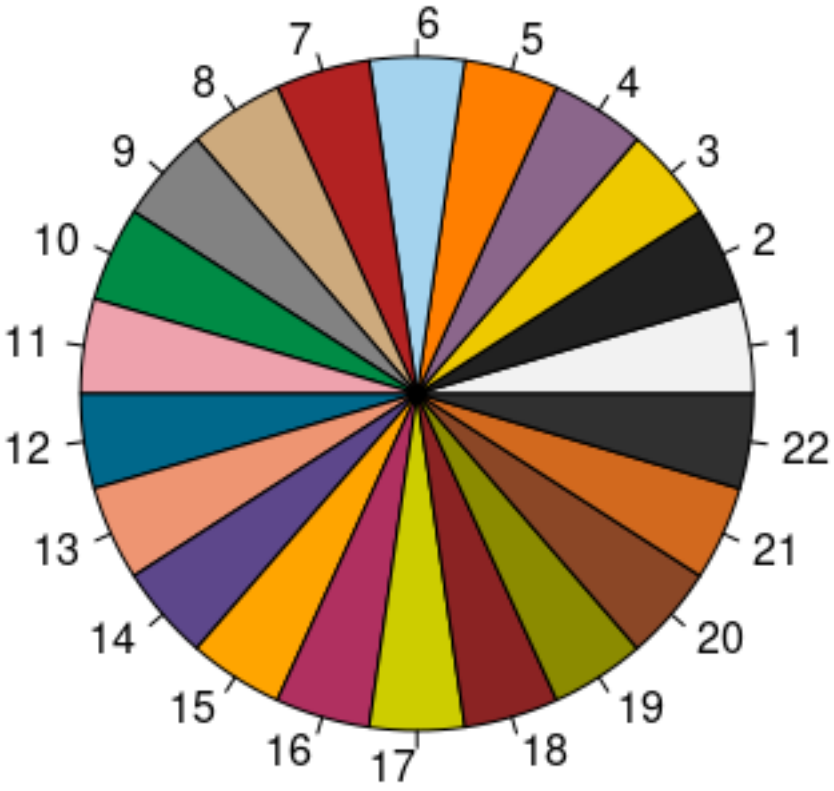
C2 R Code

C2.1 Color palette

The color palette was generated according to the high contrast palette developed by Kelly (1965).

```
kelly.colours <- c("gray95", "gray13", "gold2", "plum4", "darkorange1", "lightskyblue2", "firebrick", "burlywood3", "gray51", "springgreen4", "lightpink2", "deepskyblue4", "lightsalmon2", "mediumpurple4", "orange", "maroon", "yellow3", "brown4", "yellow4", "sienna4", "chocolate", "gray19")

pie(rep(1, length(kelly.colours)), col = kelly.colours)
```



C2.2 Code to create a table of labels for the Tukey test

To extract the labels of Tukey test anduse them in plots generated with ggplot2, the following function was used

```
generate_label_df <- function(HSD, flev, n, d){
 #Extract labels and factor levels from Tukey post-hoc
 Tukey.levels <- HSD[[flev]][,4]
 Tukey.labels <- multcompLetters(Tukey.levels)['Letters']
 plot.labels <- names(Tukey.labels[['Letters']])
```

```
 # Get highest quantile for Tukey's 5 number summary and add a bit of space to buffer
between
 # upper quantile and label placement
 boxplot.df <- ddply(d, flev, function (x) max(fivenum(x[[n]])) + 0.8)
 # Create a data frame out of the factor levels and Tukey's homogenous group letters
 plot.levels <- data.frame(lab = plot.labels,
  labels = Tukey.labels[['Letters']],
  stringsAsFactors = FALSE)
 # Merge it with the labels
 labels.df <- merge(plot.levels, boxplot.df, by.x = 'lab', by.y = flev, sort = FALSE)
 return(labels.df)
}
```

C2.3 Code for boxplots

```
library(ggplot2)
library(magrittr)
library(multcompView)
library(plyr)
library(car)

genos_monos%>%
```

```
 ggplot(aes(x=name, y = AraE))+
 geom_boxplot(aes(fill = pop_dapc8),position=position_dodge(width=0.2)) +
 geom_text(data=generate_label_df(HSD.AraE.n,'name',
which(colnames(genos_monos)=="AraE"), genos_monos), aes(x = lab, y = V1, label
= labels, colour = factor(labels)), angle = 0, family = "Arial", size = 7, color= "black")
+
 scale_y_continuous(expand = c(0.05,0.05)) +
 stat_summary(fun.y=mean, colour="black",shape = 19, geom="point",
 size=1,show_guide = FALSE) +
 geom_hline(yintercept=median(genos_monos$AraE, na.rm = T), color = "red")+
 geom_hline(yintercept=x.lower.AraE, color = "blue", linetype = "dashed")+
 geom_hline(yintercept=x.upper.AraE, color = "blue", linetype = "dashed")+
 annotate("rect", xmin = -Inf, xmax = Inf, ymin = x.lower.AraE, ymax = x.upper.AraE,
fill = "black", alpha = .1, color = NA)+
 labs(title = NULL, x = "Genotype", y = "AraE (% of Ara)", fill = "Genetic\ngroup")+
 theme_bw()+
 scale_fill_manual(values = kelly.colours)+
 my_theme.2
```

Appendix D Chapter 3

D1 Summary statistics

Table D. 1 Monosaccharides enzymatically released from the CWM

Mean and standard deviation of the monosaccharide enzymatically released from the CWM. Three monosaccharides were quantified: arabinose, glucose and xylose. The amount was expressed as a percentage (w/w) of the CWM (Ara_enz_p, Glc_Enz_p and Xyl_Enz_p) and of the total monosaccharide content (AraE, GlcE and XylE). Values were obtained from 4 technical replicates of the experiment. Values are the mean and standard deviation of the mean (sd) between 3 biological replicates (n = 3).

Genetic group	Genotype	Ara_Enz_p		Glc_Enz_p		Xyl_Enz_p		AraE		GlcE		XylE	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Flor_Taiwan	Gawain	0.087	0.012	3.206	0.86	0.439	0.127	4.725	2.984	12.06	7.484	2.108	0.719
	Geraint	0.242	0.019	4.547	1.118	0.753	0.113	7.786	0.228	10.41	3.372	2.656	0.338
	Lancelot	0.176	0.051	4.813	0.4	0.741	0.235	6.195	1.308	18.271	6.394	4.01	2.065
	Llyr	0.219	0.036	5.304	0.111	0.826	0.092	8.718	0.425	18.648	6.411	4.095	0.518

Genetic group	Genotype	Ara_Enz_p		Glc_Enz_p		Xyl_Enz_p		AraE		GlcE		XylE	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
	Percival	0.267	0.039	5.111	1.587	0.84	0.121	9.929	0.535	15.977	6.38	3.776	1.11
	Total	0.198	0.071	4.596	1.121	0.72	0.195	7.35	2.429	15.406	6.26	3.377	1.294
Hyb.gig	Cilydd	0.076	0.014	1.703	0.347	0.424	0.059	4.396	0.378	3.766	0.786	1.427	0.172
	Lamorak	0.068	0.039	0.937	0.163	0.218	0.033	3.346	0.735	2.177	0.269	0.784	0.157
	Olwen	0.154	0.005	3.573	0.636	0.461	0.034	7.235	0.633	8.416	1.921	1.541	0.104
	Total	0.099	0.046	2.071	1.232	0.368	0.12	4.712	1.707	4.786	2.996	1.251	0.376
Hyb.mx	Galahd	0.372	0.48	2.832	0.998	0.488	0.125	5.947	0.533	7.166	2.195	1.725	0.407
	Tristan	0.183	0.053	3.93	0.698	0.782	0.17	9.712	1.14	9.503	0.869	2.519	0.401
	Total	0.278	0.323	3.381	0.977	0.635	0.209	7.829	2.292	8.334	1.967	2.122	0.565
Hyb.sin	Cigfa	0.337	0.444	2.212	0.642	0.397	0.018	3.128	0.622	5.285	1.875	1.22	0.112
	Hafgan	0.115	0.063	2.524	1.568	0.453	0.183	4.358	2.072	6.197	4.313	1.481	0.764
	Nimue	0.06	0.019	2.141	0.619	0.291	0.108	3.03	1.255	5.03	1.6	0.942	0.376
	Total	0.171	0.258	2.292	0.919	0.38	0.128	3.565	1.413	5.504	2.54	1.215	0.488

Genetic group	Genotype	Ara_Enz_p		Glc_Enz_p		Xyl_Enz_p		AraE		GlcE		XylE	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Lut	Blodeuwedd	0.49	0.632	1.609	0.446	0.598	0.171	6.659	3.714	3.506	0.734	1.998	0.581
	Bran	0.096	0.015	2.331	0.305	0.687	0.046	4.731	2.519	5.42	0.699	2.258	0.15
	Eigyr	0.063	0.021	2.501	0.714	0.711	0.081	4.176	2.186	6.855	1.446	2.616	0.339
	Goronwy	0.103	0.013	1.687	0.396	0.738	0.164	5.412	2.473	4.116	0.855	2.253	0.38
	Govannon	0.342	0.459	1.39	0.378	0.492	0.089	3.86	1.156	3.355	0.94	1.669	0.242
	Gwalchmei	0.103	0.029	2.525	0.575	0.942	0.33	7.173	1.027	5.94	1.161	2.948	1.064
	Gwydion	0.085	0.017	2.171	0.72	0.576	0.135	4.676	1.984	5.192	1.769	1.903	0.47
	Kay	0.392	0.451	2.791	0.426	1.132	0.155	7.497	0.696	6.959	1.522	3.581	0.537
	Lleu	0.049	0.004	1.218	0.092	0.448	0.043	2.681	0.613	3.02	0.636	1.541	0.128
	Mabon	0.099	0.008	1.64	0.352	0.626	0.127	5.351	1.639	4.013	1.179	2.149	0.499
	Math	0.117	0.008	2.326	0.401	0.792	0.088	6.601	2.774	5.62	0.974	2.536	0.147
	Owain	0.097	0.025	2.892	0.38	0.833	0.045	4.177	1.13	7.014	0.845	2.537	0.188
	Total	0.194	0.318	2.053	0.664	0.706	0.221	5.266	2.404	4.963	1.69	2.307	0.671

Genetic group	Genotype	Ara_Enz_p		Glc_Enz_p		Xyl_Enz_p		AraE		GlcE		XylE	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Sacc_rob	Aranrhod	0.521	0.729	2.109	0.52	0.692	0.198	5.437	2.225	4.617	0.924	2.248	0.638
	Beli	0.068	0.011	1.25	0.33	0.557	0.061	5.231	1.476	5.323	3.791	2.821	1.521
	Culhwch	0.083	0.016	2.58	0.398	0.728	0.151	4.266	2.561	5.532	1.044	2.237	0.658
	Gaheris	0.054	0.011	2.349	0.642	0.446	0.05	3.647	0.048	5.042	1.264	1.523	0.162
	Gwenhwyfar	0.085	0.019	2.093	0.388	0.769	0.194	4.747	1.952	4.688	0.662	2.41	0.531
	Lunete	0.104	0.011	1.898	0.608	0.701	0.212	6.027	0.318	6.596	5.046	2.995	1.731
	Modred	0.081	0.019	1.783	0.259	0.68	0.07	6.278	2.576	7.49	4.543	3.705	1.616
	Nudd	0.126	0.027	2.948	0.82	0.741	0.157	5.049	2.685	9.549	2.683	3.15	1.092
	Pryderi	0.467	0.563	3.187	1.979	0.854	0.214	10.865	8.12	11.956	4.509	4.773	2.767
	Pwyll	0.154	0.054	2.721	1.361	0.713	0.096	13.128	8.096	8.212	6.737	2.924	1.272
	Rhiannon	0.276	0.269	3.434	0.771	0.805	0.09	5.455	1.889	8.144	2.178	2.53	0.364
	Total	0.187	0.293	2.431	0.958	0.703	0.165	6.415	4.391	7.066	3.721	2.848	1.38
Sin (EMI/PRI)	Luned	0.071	0.021	1.984	0.271	0.376	0.085	2.35	0.903	5.032	0.579	1.186	0.328

Genetic group	Genotype	Ara_Enz_p		Glc_Enz_p		Xyl_Enz_p		AraE		GlcE		XylE	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
	Total	0.071	0.021	1.984	0.271	0.376	0.085	2.35	0.903	5.032	0.579	1.186	0.328
Sin_N_Japan	Aeron	0.154	0.072	3.852	1.396	0.538	0.16	4.751	2.442	10.108	3.699	1.924	0.624
	Dylan	0.059	0.011	1.927	0.552	0.451	0.083	2.386	0.129	6.319	2.117	1.701	0.255
	Enid	0.101	0.013	2.945	0.279	0.345	0.113	4.372	0.492	10.021	4.37	1.702	1.149
	Gareth	0.236	0.023	4.968	0.67	0.664	0.189	6.148	0.5	13.124	1.142	2.242	0.762
	Peredur	0.163	0.137	1.976	0.613	0.491	0.103	9.421	7.251	8.976	7.886	2.128	0.551
	Total	0.142	0.086	3.134	1.374	0.498	0.158	5.44	4.06	9.71	4.419	1.918	0.642
Sin_S_Japan	Angharad	0.329	0.156	3.558	1.9	0.544	0.259	10.495	4.282	11.382	2.863	2.153	0.511
	Bedivere	0.377	0.409	1.77	0.161	0.274	0.083	2.804	0.295	4.439	0.344	0.904	0.284
	Urien	0.154	0.054	3.016	0.695	0.482	0.187	4.864	2.533	7.727	1.323	1.663	0.642
	Total	0.287	0.243	2.781	1.289	0.433	0.206	6.46	4.356	7.849	3.4	1.573	0.697
Sin_Taiwan	Bors	0.115	NA	1.15	NA	0.339	NA	NA	NA	NA	NA	NA	NA
	Bors	0.047	0.002	1.412	0.095	0.343	0.096	1.957	0.342	3.807	0.513	1.107	0.239

Genetic group	Genotype	Ara_Enz_p		Glc_Enz_p		Xyl_Enz_p		AraE		GlcE		XylE	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
	Uther	0.251	0.314	2.029	0.559	0.483	0.129	2.648	0.792	5.54	1.321	1.693	0.375
	Yspaddaden	0.48	0.685	2.609	0.42	0.457	0.058	3.993	1.08	6.051	1.668	1.375	0.276
	Total	0.26	0.421	2.017	0.627	0.428	0.107	2.866	1.132	5.133	1.495	1.392	0.365
		0.191	0.262	2.62	1.225	0.601	0.223	5.54	3.34	7.28	4.525	2.225	1.112

Table D. 2 Summary statistics of the cell wall composition variables

Mean and standard deviation of the monosaccharide content in the CWM. Five monosaccharides were quantified: arabinose (Ara), galactose (Gal), glucose (Glc), xylose (Xyl) and mannose (Man). Values were obtained from 4 technical replicates of the experiment. Values are the mean and standard deviation of the mean (sd) between 3 biological replicates (n = 3).

Genetic group	Genotype	Ara		Gal		Glc		Xyl		Man	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Flor_Taiwan	Gawain	2.221	0.833	0.696	0.309	32.241	14.084	22.215	7.578	1.673	0.703

Genetic group	Genotype	Ara		Gal		Glc		Xyl		Man	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
	Geraint	3.193	0.277	1.085	0.178	38.867	0.051	26.036	0.554	1.514	0.904
	Lancelot	2.998	1.105	0.699	0.217	29.635	7.664	20.86	3.932	5.93	7.993
	Llyr	2.574	0.638	0.756	0.203	30.818	10.212	20.156	5.493	6.358	8.612
	Percival	2.854	0.209	0.986	0.228	35.991	4.513	24.69	1.902	1.456	0.271
	Total	2.75	0.702	0.833	0.255	33.128	8.362	22.56	4.623	3.52	5.194
Hyb.gig	Cilydd	1.717	0.188	0.456	0.038	45.564	1.056	29.92	3.71	20.38	31.815
	Lamorak	1.906	0.797	0.619	0.424	42.781	3.102	28.061	2.142	1.537	0.469
	Olwen	2.226	0.171	0.73	0.093	42.669	3.156	30.007	1.221	3.799	3.192
	Total	1.915	0.49	0.586	0.258	43.671	2.682	29.33	2.422	8.572	18.304
Hyb.mx	Galahd	1.582	0.124	0.381	0.084	40.051	7.369	28.674	5.601	6.326	8.023
	Tristan	2.212	0.229	0.609	0.067	41.469	3.39	31.386	2.982	7.867	10.97
Hyb.sin	Cigfa	2.183	0.208	0.463	0.113	42.337	2.579	32.468	2.668	19.217	30.323
	Hafgan	2.531	0.509	0.551	0.309	42.596	4.209	31.809	3.607	18.476	30.204

Genetic group	Genotype	Ara		Gal		Glc		Xyl		Man	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
	Nimue	2.155	0.067	0.585	0.026	42.942	1.404	30.99	0.707	18.255	30.318
	Total	1.897	0.394	0.495	0.146	40.76	5.189	30.03	4.279	7.096	8.637
Lut	Blodeuwedd	2.059	0.892	0.925	0.699	44.017	0.445	30.04	0.801	2.026	0.64
	Bran	2.343	0.967	1.033	0.654	43.146	1.876	30.432	0.595	2.56	0.516
	Eigyr	1.976	1.061	1.152	0.961	37.1	8.697	27.734	6.187	3.508	2.659
	Goronwy	2.3	1.324	1.032	1.091	40.79	1.417	32.611	3.501	3.318	3.279
	Govannon	2.246	0.892	0.873	0.676	41.597	0.414	29.497	1.783	2.071	1.251
	Gwalchmei	1.715	0.038	0.613	0.191	42.37	2.13	32.048	2.112	1.925	1.767
	Gwydion	2.02	0.74	0.703	0.507	42.079	1.19	30.378	1.052	2.13	0.686
	Kay	1.761	0.019	0.505	0.011	40.882	3.465	31.765	1.601	2.491	0.899
	Lleu	2.046	0.796	0.687	0.482	41.149	4.814	29.293	3.499	1.721	0.604
	Mabon	2.112	0.935	0.817	0.613	42.062	5.375	29.967	1.222	2.569	0.224
	Math	2.141	1.148	0.798	0.833	41.495	0.557	31.281	1.803	1.897	1.374

Genetic group	Genotype	Ara		Gal		Glc		Xyl		Man	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
	Owain	2.594	1.537	1.32	1.376	41.361	1.93	32.895	2.696	4.509	4.876
	Total	2.119	0.851	0.871	0.667	41.697	3.332	30.614	2.578	2.519	1.819
Sacc_rob	Aranrhod	2.218	1.252	1.053	1.236	45.386	4.196	30.966	2.492	3.315	2.32
	Beli	1.618	0.363	0.674	0.004	33.894	6.831	25.967	5.284	1.566	0.399
	Culhwch	2.716	1.892	1.379	1.634	46.926	1.673	33.205	3.475	6.75	7.683
	Gaheris	1.474	0.298	0.349	0.046	46.51	3.112	29.394	0.67	25.075	41.002
	Gwenhwyfar	2.052	0.897	0.779	0.683	44.537	3.083	31.901	2.771	4.437	3.182
	Lunete	1.721	0.16	0.552	0.243	36.258	10.118	25.949	5.462	1.398	0.511
	Modred	1.131	0.347	1.889	2.569	31.4	11.949	21.894	6.305	5.816	8.093
	Nudd	2.725	0.895	0.602	0.281	31.771	7.559	24.83	5.708	7.053	8.683
	Pryderi	1.932	1.337	0.664	0.486	30.66	16.493	23.695	11.998	2.226	1.753
	Pwyll	1.467	0.101	0.428	0.088	40.132	10.683	27.015	6.43	1.759	0.4
	Rhiannon	2.356	0.656	0.768	0.461	43.028	2.8	31.967	2.377	2.089	1.804

Genetic group	Genotype	Ara		Gal		Glc		Xyl		Man	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
	Total	1.957	0.927	0.835	0.981	39.301	9.424	27.949	5.942	5.715	12.921
Sin (EMI/PRI)	Luned	3.09	0.481	0.823	0.133	39.578	2.861	32.236	2.328	1.357	0.688
	Sin (EMI/PRI)	3.09	0.481	0.823	0.133	39.578	2.861	32.236	2.328	1.357	0.688
Sin_N_Japan	Aeron	3.64	0.917	1.127	0.093	38.11	2.332	28.098	1.368	1.767	0.555
	Dylan	2.612	0.265	0.715	0.162	33.147	8.971	27.876	6.464	1.627	0.471
	Enid	2.263	0.171	0.538	0.179	33.692	9.509	25.797	7.422	2.136	0.826
	Gareth	3.779	0.2	1.06	0.049	37.788	1.739	29.649	1.839	1.27	0.423
	Peredur	1.84	0.524	0.387	0.137	31.845	16.439	24.603	8.704	1.886	1.108
	Total	2.797	0.909	0.791	0.313	34.916	8.438	27.03	5.505	1.77	0.683
Sin_S_Japan	Angharad	3.281	0.302	1.125	0.251	32.061	6.026	26.971	3.174	1.904	0.311
	Bedivere	2.385	0.329	0.657	0.111	40.369	2.388	27.47	4.717	3.968	4.403

Genetic group	Genotype	Ara		Gal		Glc		Xyl		Man	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
	Urien	3.537	0.976	1.523	0.796	38.763	2.611	28.996	0.907	3.731	2.76
	Total	3.153	0.743	1.157	0.574	37.064	5.174	27.812	3.02	3.201	2.781
Sin_Taiwan	Bors	2.411	0.4	0.571	0.152	37.334	4.057	30.655	2.46	1.438	0.195
	Uther	2.338	0.364	0.624	0.207	38.219	1.026	29.774	1.039	1.069	0.366
	Yspaddaden	2.628	0.388	0.581	0.237	44.169	6.081	33.549	3.257	19.582	31.458
	Total	2.459	0.358	0.592	0.176	39.907	4.898	31.326	2.713	7.363	18.206

Table D. 3 Cell wall architecture: proportion between monosaccharides

Mean and standard deviation of the ratios between monosaccharides the CWM.The ratios were selected because they jhabe been linked to cell wall recalcitrance. Values for the ratio between cellulose and hemicellulose (Cell/Hem), xylose and arabinose (Xyl/Ara), mannose and

galactose (Man/Gal) ,xylose and glucose (Xyl/Glc) are presented here. Values were obtained from 4 technical replicates of the experiment. Values are the mean and standard deviation of the mean (sd) between 3 biological replicates (n = 3).

Genetic group	Genotype	Cell/Hem		Xyl/Ara		Man/Gal		Xyl/Glc	
		mean	sd	mean	sd	mean	sd	mean	sd
Flor_Taiwan	Gawain	1.368	0.152	10.118	0.647	2.457	0.284	0.718	0.102
	Geraint	0.896	0.058	8.193	0.884	1.345	0.612	0.67	0.013
	Lancelot	1.126	0.061	7.518	2.463	10.74	15.857	0.712	0.046
	Llyr	1.11	0.104	7.713	1.105	1.699	1.283	0.665	0.053
	Percival	1.008	0.088	8.693	1.04	1.492	0.18	0.689	0.035
	Total	1.102	8.523	3.858	0.692	0.182	1.566	7.593	0.055
Hyb.gig	Cilydd	1.834	0.221	17.405	0.415	47.319	74.536	0.658	0.09
	Lamorak	1.914	0.134	16.01	4.63	3.179	1.879	0.657	0.044
	Olwen	1.537	0.162	13.767	1.593	2.673	0.873	0.704	0.023
	Total	1.762	15.972	19.605	0.673	0.229	2.968	45.991	0.057
Hyb.mx	Galahd	1.562	0.166	20.18	3.281	4.806	3.021	0.716	0.053

Genetic group	Genotype	Cell/Hem		Xyl/Ara		Man/Gal		Xyl/Glc	
		mean	sd	mean	sd	mean	sd	mean	sd
	Tristan	1.313	0.114	15.049	2.043	2.507	0.239	0.757	0.026
	Total	1.437	17.615	3.656	0.736	0.186	3.709	2.196	0.043
Hyb.sin	Cigfa	1.588	0.066	14.898	0.805	52.86	85.934	0.771	0.11
	Hafgan	1.401	0.17	13.055	3.651	69.167	117.327	0.747	0.04
	Nimue	1.574	0.085	14.215	0.223	1.292	0.348	0.722	0.021
	Total	1.521	14.036	46.083	0.747	0.135	2.177	82.85	0.063
Lut	Blodeuwedd	1.46	0.546	16.564	5.514	2.658	1.404	0.683	0.023
	Bran	1.583	0.101	14.081	5.463	3.068	1.355	0.706	0.04
	Eigyr	1.581	0.18	15.761	5.1	3.815	0.317	0.749	0.024
	Goronwy	1.461	0.154	16.601	6.353	3.616	1.476	0.798	0.06
	Govannon	1.702	0.25	14.301	4.422	2.83	1.714	0.709	0.047
	Gwalchmei	1.681	0.105	18.799	2.145	2.784	1.72	0.756	0.022
	Gwydion	1.674	0.324	16.251	4.988	3.888	1.807	0.722	0.014

Genetic group	Genotype	Cell/Hem		Xyl/Ara		Man/Gal		Xyl/Glc	
		mean	sd	mean	sd	mean	sd	mean	sd
	Kay	1.403	0.098	17.643	1.01	3.917	0.46	0.779	0.05
	Lleu	1.783	0.149	15.243	3.612	2.905	0.998	0.713	0.046
	Mabon	1.788	0.085	15.959	6.084	4.132	1.984	0.718	0.072
	Math	1.693	0.099	16.929	6.544	2.986	1.061	0.754	0.04
	Owain	1.533	0.176	15.026	5.949	3.206	1.128	0.796	0.073
	Total	1.6	16.073	3.273	0.736	0.269	4.514	1.28	0.054
Sacc_rob	Aranrhod	1.776	0.244	16.524	6.995	4.688	2.004	0.684	0.055
	Beli	1.686	0.265	16.085	0.346	2.322	0.578	0.766	0.002
	Culhwch	1.643	0.073	15.413	6.979	5.076	1.555	0.708	0.072
	Gaheris	1.835	0.055	20.491	4.199	65.346	105.762	0.633	0.03
	Gwenhwyfar	1.639	0.191	17.092	5.378	6.244	0.996	0.719	0.084
	Lunete	1.867	0.135	15.35	4.374	2.803	1.392	0.727	0.066
	Modred	1.844	0.109	20.515	7.728	3.236	1.837	0.717	0.085

Genetic group	Genotype	Cell/Hem		Xyl/Ara		Man/Gal		Xyl/Glc	
		mean	sd	mean	sd	mean	sd	mean	sd
	Nudd	1.302	0.186	9.806	3.516	21.932	33.38	0.785	0.069
	Pryderi	1.45	0.291	13.773	4.207	3.27	0.175	0.789	0.048
	Pwyll	1.763	0.104	18.583	5.224	4.373	1.905	0.677	0.025
	Rhiannon	1.462	0.083	14.103	2.873	2.555	0.869	0.746	0.086
	Total	1.661	16.16	11.35	0.722	0.232	5.285	33.697	0.07
Sin (EMI/PRI)	Luned	1.366	0.079	10.536	1.03	1.777	1.222	0.815	0.035
	Total	1.366	10.536	1.777	0.815	0.079	1.03	1.222	0.035
Sin_N_Japan	Aeron	1.169	0.136	8.006	1.712	1.587	0.569	0.738	0.014
	Dylan	1.279	0.093	10.747	2.797	2.272	0.305	0.848	0.045
	Enid	1.489	0.065	13.328	1.984	4.865	0.625	0.764	0.053
	Gareth	1.034	0.071	7.843	0.072	1.209	0.455	0.779	0
	Peredur	1.483	0.122	14.551	8.31	3.46	1.33	0.852	0.232

Genetic group	Genotype	Cell/Hem		Xyl/Ara		Man/Gal		Xyl/Glc	
		mean	sd	mean	sd	mean	sd	mean	sd
	Total	1.291	10.943	2.554	0.798	0.202	4.648	1.42	0.107
Sin_S_Japan	Angharad	1.184	0.186	8.24	0.916	1.707	0.107	0.849	0.069
	Bedivere	1.538	0.095	12.788	1.837	2.415	2.129	0.678	0.088
	Urien	1.322	0.219	8.623	2.322	2.588	1.28	0.751	0.064
	Total	1.348	9.521	2.214	0.759	0.216	2.522	1.14	0.098
Sin_Taiwan	Bors	1.41	NA	NA	NA	NA	NA	NA	NA
	Bors	1.399	0.053	12.989	2.782	2.695	0.988	0.823	0.024
	Uther	1.457	0.116	12.95	2.085	1.883	1.021	0.779	0.007
	Yspaddaden	1.283	0.073	13.02	2.789	60.387	101.084	0.762	0.029
	Total	1.383	12.986	21.655	0.788	0.1	2.229	58.301	0.033

Table D. 4 Summary statistics of leaf to stem ratio (LSR)

The leaf to stem ratio values (LSR) were calculated as a proportion between dry weight of the leaf (DWL) and dry weight of the stem (DWS) in g. Values are the mean and the standard deviation of the mean (sd)

		DWL (g)		DWS (g)		LSR	
Genetic group	Genotype	mean	sd	mean	sd	mean	sd
Flor_Taiwan	Gawain	48.177	10.449	8.9	7.978	7.707	3.924
	Geraint	36.957	13.512	12.8	2.232	3.033	1.445
	Lancelot	18.63	8.429	16.227	15.429	2.573	2.916
	Llyr	30.39	3.939	10.64	3.77	3	0.622
	Percival	7.72	2.643	3.71	2.555	2.427	0.805
	Total	28.375	16.355	10.455	8.085	3.748	2.849
Hyb.gig	Cilydd	17.19	5.305	79.413	24.538	0.227	0.059
	Lamorak	15.647	7.308	64.193	3.824	0.247	0.112
	Olwen	13.215	7.799	31.725	33.482	0.65	0.438

Genetic group	Genotype	DWL (g)		DWS (g)		LSR	
		mean	sd	mean	sd	mean	sd
	Total	15.617	5.891	61.784	27.022	0.34	0.262
Hyb.mx	Galahd	11.07	7.548	29.977	14.47	0.35	0.07
	Tristan	19.543	1.014	32.017	3.159	0.613	0.072
	Total	15.307	6.689	30.997	9.434	0.482	0.158
Hyb.sin	Cigfa	14.11	3.68	28.477	6.255	0.497	0.07
	Hafgan	7.38	2.128	15.453	1.956	0.47	0.075
	Nimue	15.793	2.151	41.183	7.168	0.387	0.015
	Total	12.428	4.532	28.371	12.154	0.451	0.072
Lut	Blodeuwedd	4.69	1.199	204.47	16.945	0.023	0.01
	Bran	0.28	NA	38.54	NA	0.01	NA
	Eigyr	15.32	NA	8.43	NA	1.82	NA
	Goronwy	7.885	5.226	125.44	95.431	0.065	0.007
	Govannon	4.285	0.686	233.265	12.325	0.02	0

Genetic group	Genotype	DWL (g)		DWS (g)		LSR	
		mean	sd	mean	sd	mean	sd
	Gwalchmei	0.8	1.131	46.29	65.464	0.02	NA
	Gwydion	4.495	3.84	59.475	14.39	0.07	0.042
	Kay	2.28	0.651	39.123	16.712	0.067	0.031
	Lleu	1.935	2.454	60.38	18.54	0.04	0.057
	Mabon	5.06	1.045	137.6	32.647	0.04	0.01
	Math	1.603	1.983	168.397	53.053	0.007	0.012
	Owain	3.94	1.042	27.93	4.34	0.14	0.026
	Total	4.04	3.239	114.755	80.726	0.108	0.332
Sacc_rob	Aranrhod	1.937	1.678	21.423	18.891	0.09	0.014
	Beli	1.075	1.52	37.255	52.687	0.03	NA
	Culhwch	1.857	1.625	23.297	20.185	0.08	0.014
	Gaheris	1.747	1.68	13.593	12.947	0.13	0
	Gwenhwyfar	4.16	4.244	31.47	12.791	0.11	0.08

Genetic group	Genotype	DWL (g)		DWS (g)		LSR	
		mean	sd	mean	sd	mean	sd
	Lunete	4.815	1.93	100.597	49.093	0.05	0.042
	Modred	2.33	0.304	58.347	5.187	0.04	0
	Nudd	14.735	9.935	66.005	7.474	0.22	0.127
	Pryderi	2.613	1.881	21.173	15.98	0.127	0.012
	Pwyll	4.393	1.734	16.817	5.021	0.26	0.07
	Rhiannon	4.227	1.243	18.737	7.628	0.25	0.111
	Total	3.701	4.054	36.222	32.527	0.136	0.096
Sin (EMI/PRI)	Luned	6.013	1.435	12.497	6.076	0.513	0.108
	Total	6.013	1.435	12.497	6.076	0.513	0.108
Sin_N_Japan	Aeron	31.95	3.026	47.055	9.412	0.695	0.205
	Dylan	11.23	8.782	11.15	5.049	0.92	0.368
	Enid	4.74	1.753	11.66	5.49	0.443	0.191

Genetic group	Genotype	DWL (g)		DWS (g)		LSR	
		mean	sd	mean	sd	mean	sd
	Gareth	34.51	6.46	32.613	8.399	1.153	0.537
	Peredur	8.217	0.769	21.883	11.049	0.433	0.18
	Total	17.597	13.809	24.222	14.847	0.717	0.409
Sin_S_Japan	Angharad	4.863	1.695	7.627	1.276	0.667	0.289
	Bedivere	14.54	3.532	28.437	8.093	0.517	0.055
	Urien	22.817	0.89	34.643	2.434	0.66	0.026
	Total	14.073	8.037	23.569	12.979	0.614	0.165
Sin_Taiwan	Bors	12.21	2.469	24.753	13.989	0.637	0.376
	Uther	12.57	0.127	18.355	3.486	0.7	0.141
	Yspaddaden	5.337	0.933	7.36	1.853	0.773	0.285
	Total	9.723	3.899	16.631	11.162	0.704	0.265

D1.1 ANOVA tables

Table D. 5 ANOVA test for the effect of genotype and genetic group on the enzymatic monosaccharide release

Statistical test of the effect of genotype and genetic group (tested independently) on the amount of arabinose (AraE), glucose (GlcE) and xylose (XylE) enzymatically released form the cell wall.

Variable	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power
AraE	Genotype	47	806.351	17.156	2.188	0.001	0.548	0.548	0.296	0.296	0.297	1.1	1
	Residuals	85	666.382	7.84									
	Genetic group	9	228.521	25.391	2.51	0.011	0.155	0.155	0.093	0.093	0.093	0.429	0.942
	Residuals	123	1244.212	10.116									
GlcE	Genotype	47	1977.42	42.073	4.201	0.000	0.671	0.671	0.509	0.509	0.511	1.427	1
	Residuals	97	971.401	10.014									
	Genetic group	9	1374.552	152.728	13.097	0.000	0.466	0.466	0.429	0.429	0.431	0.934	1
	Residuals	135	1574.269	11.661									
XylE	Genotype	47	111.533	2.373	3.491	0.000	0.631	0.631	0.448	0.448	0.45	1.307	1
	Residuals	96	65.257	0.68									

Variable	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power
	Genetic group	9	63.683	7.076	8.383	0.000	0.36	0.36	0.316	0.316	0.317	0.75	1
	Residuals	134	113.108	0.844									

Table D. 6 ANOVA test for the effect of genotype and genetic group on the total amount of monosaccharides in the cell wall.

Statistical test of the effect of genotype and genetic group (tested independently) on the amount of monosaccharides in the cell wall of the 49 genotypes under study.

Variable	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power
Ara	Genotype	47	42	0.894	1.404	0.087	0.434	0.434	0.124	0.124	0.125	0.876	1
	Residuals	86	54.731	0.636									
	Genetic group	9	20.536	2.282	3.713	0	0.212	0.212	0.154	0.154	0.155	0.519	0.994

Variable	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power
	Residuals	124	76.195	0.614									
Gal	Genotype	47	13.049	0.278	0.581	0.978	0.247	0.247	-0.177	-0.177	-0.179	0.573	0.834
	Residuals	83	39.678	0.478									
	Genetic group	9	2.978	0.331	0.805	0.613	0.056	0.056	-0.014	-0.014	-0.014	0.245	0.416
	Residuals	121	49.749	0.411									
Glc	Genotype	47	3185.478	67.776	1.668	0.018	0.447	0.447	0.178	0.178	0.179	0.899	1
	Residuals	97	3941.531	40.634									
	Genetic group	9	1378.306	153.145	3.596	0	0.193	0.193	0.139	0.139	0.14	0.49	0.992
	Residuals	135	5748.703	42.583									
Xyl	Genotype	47	1593.551	33.905	1.934	0.003	0.486	0.486	0.234	0.234	0.235	0.973	1
	Residuals	96	1683.234	17.534									

Variable	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power
	Genetic group	9	932.134	103.57	5.919	0	0.284	0.284	0.235	0.235	0.236	0.631	1
	Residuals	134	2344.651	17.497									
Man	Genotype	47	5152.288	109.623	0.744	0.869	0.267	0.267	-0.091	-0.091	-0.092	0.604	0.932
	Residuals	96	14143.508	147.328									
	Genetic group	9	2366.905	262.989	2.082	0.035	0.123	0.123	0.063	0.063	0.064	0.374	0.883
	Residuals	134	16928.891	126.335									

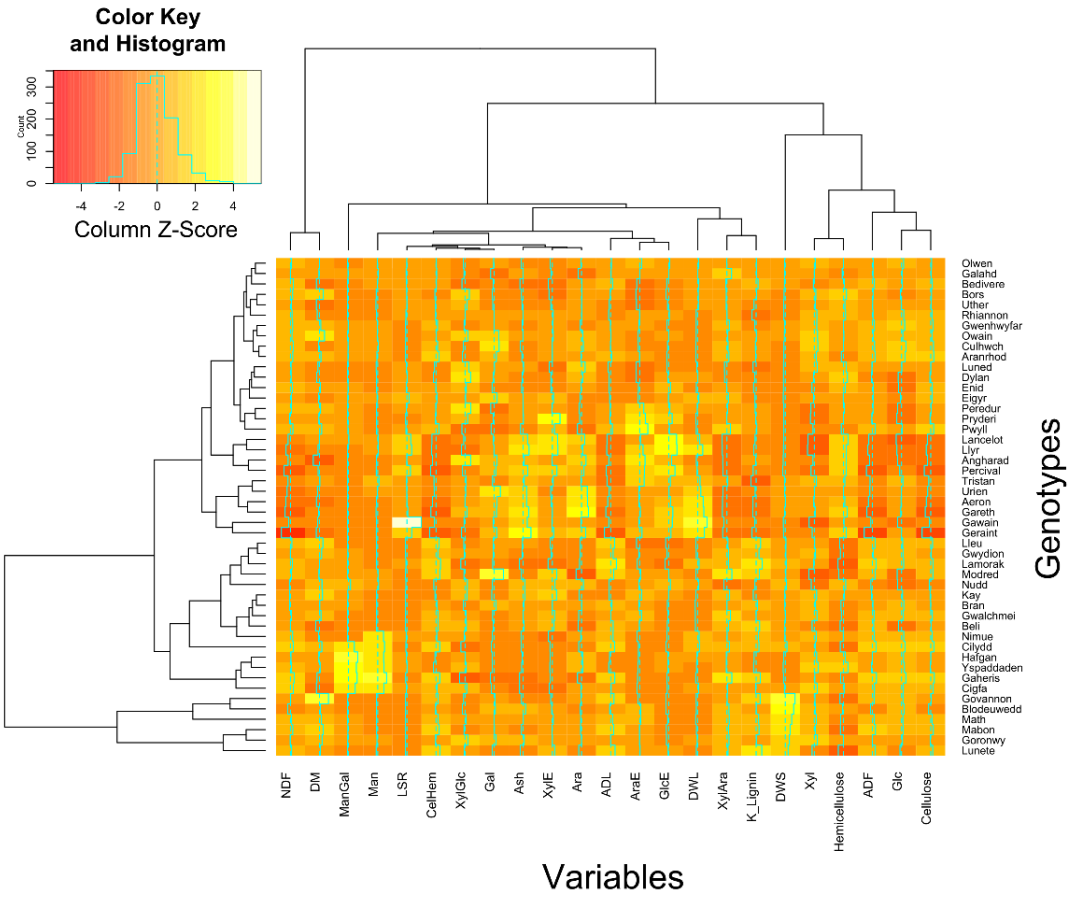
Table D. 7 ANOVA test for the effect of genotype and genetic group on the cell wall architecture.

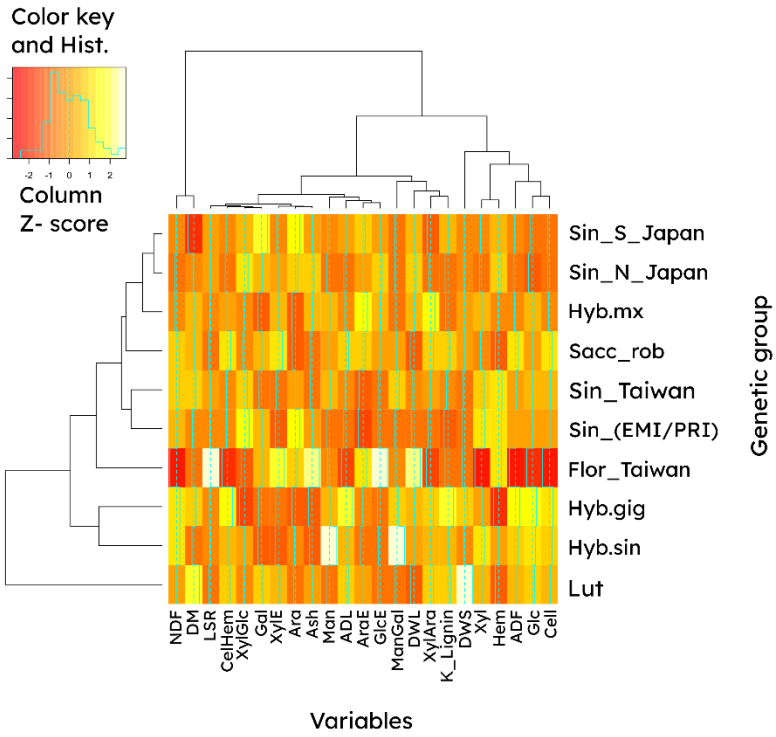
Statistical test of the effect of genotype and genetic group (tested independently) on cellulose to hemicellulose (Cell/Hem), xylose to arabinose (Xyl/Ara), xylose to glucose (Xyl/Glc) and mannose to galactose (Man/Gal) ratios in the cell wall. These variables relate to the proportion between cell wall components and monosaccharide proportion iin the carbohydrate polymers of the cell wall. They were used to describe the cell wall architecture.

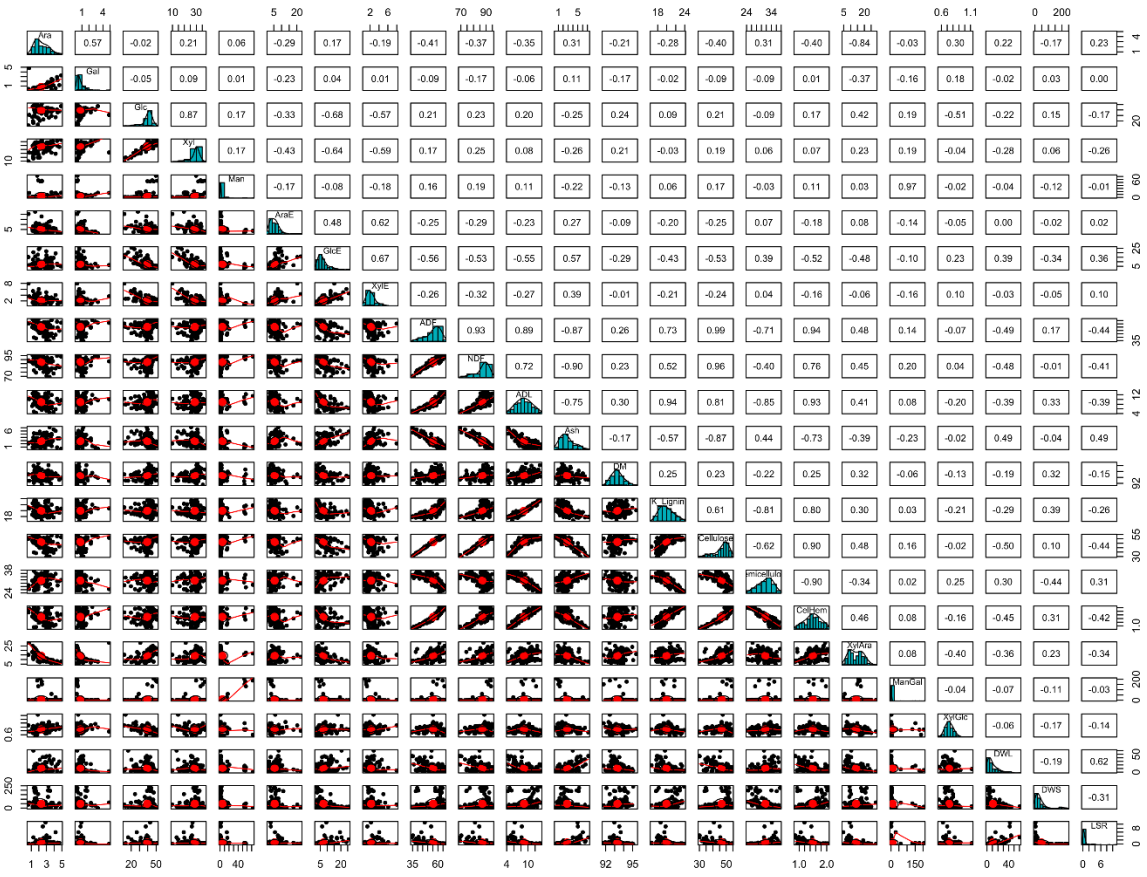
Variable	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power
CelHem	Genorype	47	8.433	0.179	4.97	0	0.7	0.7	0.558	0.558	0.559	1.528	1
	Residuals	100	3.61	0.036									
	Genetic group	9	5.319	0.591	12.132	0	0.442	0.442	0.404	0.404	0.405	0.889	1
	Residuals	138	6.723	0.049									
XylAra	Genorype	47	1624.021	34.554	1.893	0.005	0.508	0.508	0.238	0.238	0.24	1.017	1
	Residuals	86	1570.141	18.257									
	Genetic group	9	1104.23	122.692	7.28	0	0.346	0.346	0.297	0.297	0.298	0.727	1
	Residuals	124	2089.932	16.854									
XylGlc	Genorype	47	0.391	0.008	2.075	0.001	0.504	0.504	0.26	0.26	0.261	1.008	1
	Residuals	96	0.385	0.004									
	Genetic group	9	0.171	0.019	4.206	0	0.22	0.22	0.167	0.167	0.168	0.532	0.998

Variable	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power
	Residuals	134	0.605	0.005									
ManGal	Genorype	47	42952.364	913.88	0.766	0.84	0.302	0.302	-0.092	-0.092	-0.093	0.658	0.945
	Residuals	83	99067.814	1193.588									
	Genetic group	9	15977.91	1775.323	1.704	0.095	0.113	0.113	0.046	0.046	0.046	0.356	0.795
	Residuals	121	126042.268	1041.672									

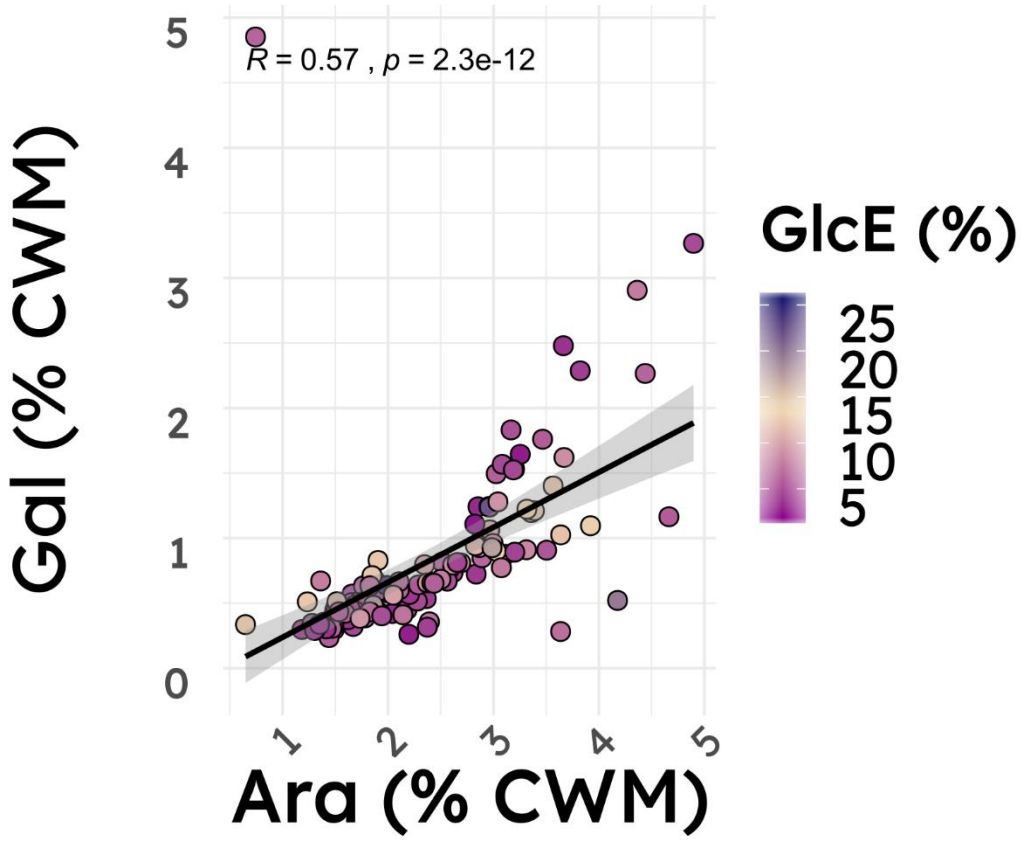
D1.2 Figures

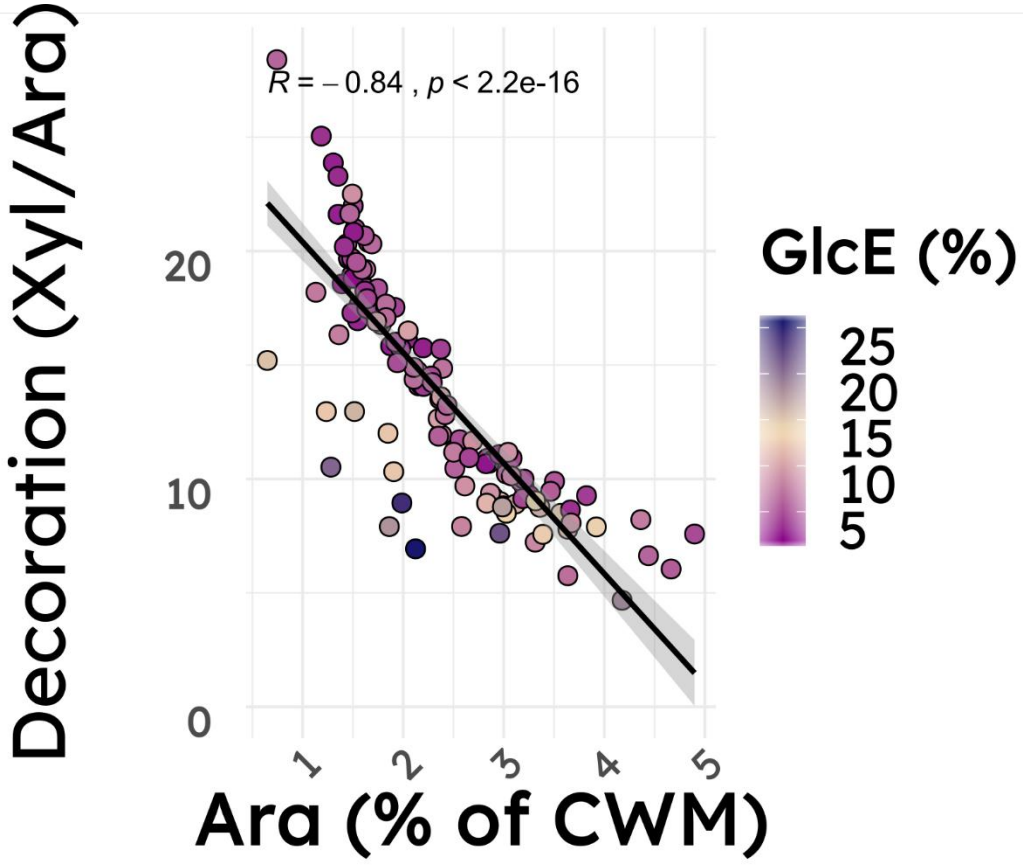

Color Key
and Histogram
Count
0
100
200
300
-4
-2
0
2
4
Column Z-Score
Olwen
Galahd
Bedivere
Bors
Uther
Rhiannon
Gwenhwyfar
Owain
Culhwch
Aranrhod
Luned
Dylan
Enid
Eigyr
Peredur
Pryderi
Pwyll
Lancelot
Llyr
Angharad
Percival
Tristan
Urien
Aeron
Gareth
Gawain
Geraint
Lleu
Gwydion
Lamorak
Modred
Nudd
Kay
Bran
Gwalchmei
Beli
Nimue
Cilydd
Hafgan
Yspaddaden
Gaheris
Cigfa
Govannon
Blodeuwedd
Math
Mabon
Goronwy
Lunete
Genotypes
NDF
DM
ManGal
Man
LSR
CelHem
XylGlc
Gal
Ash
XylE
Ara
ADL
AraE
GlcE
DWL
XylAra
K_Lignin
DWS
Xyl
Hemicellulose
ADF
Glc
Cellulose
Variables

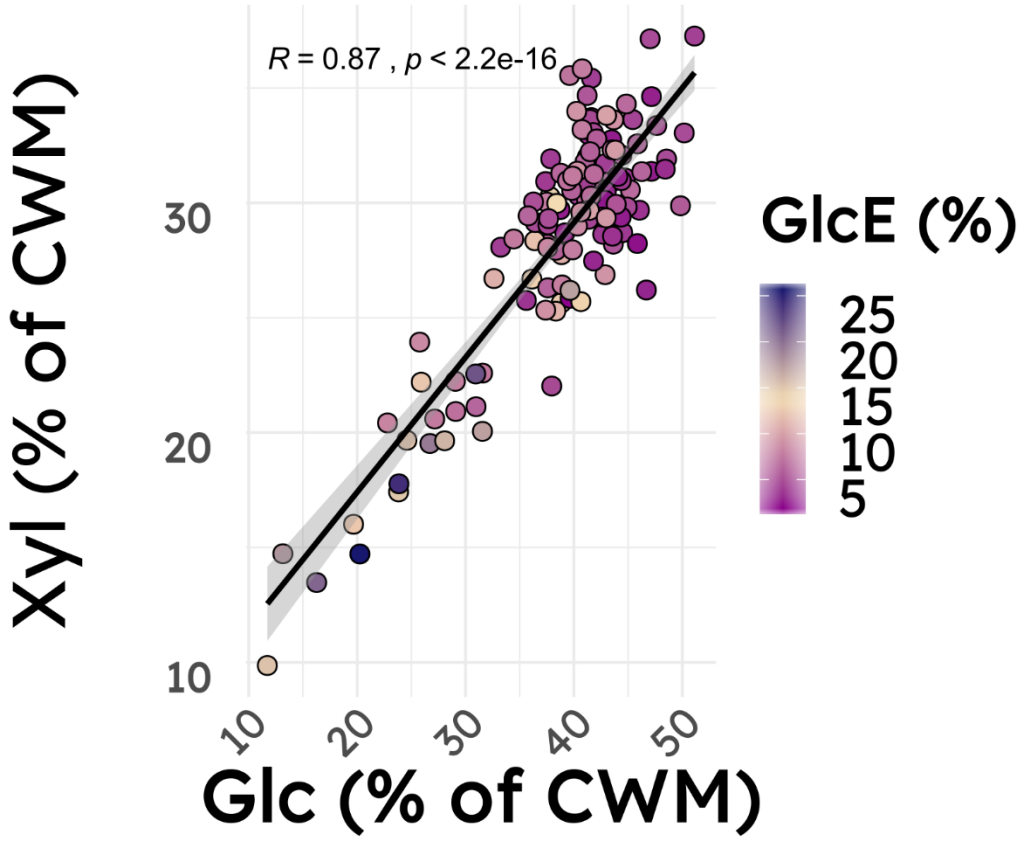

Color key
and Hist.
-2 -1 0 1 2
Column
Z- score
Sin_S_Japan
Sin_N_Japan
Hyb.mx
Sacc_rob
Sin_Taiwan
Sin_(EMI/PRI)
Flor_Taiwan
Hyb.gig
Hyb.sin
Lut
Genetic group
NDF
DM
LSR
CelHem
XylGlc
Gal
XylE
Ara
Ash
Man
ADL
AraE
GlcE
ManGal
DwL
XylAra
K_Lignin
DwS
Xyl
Hem
ADF
Glc
Cell
Variables

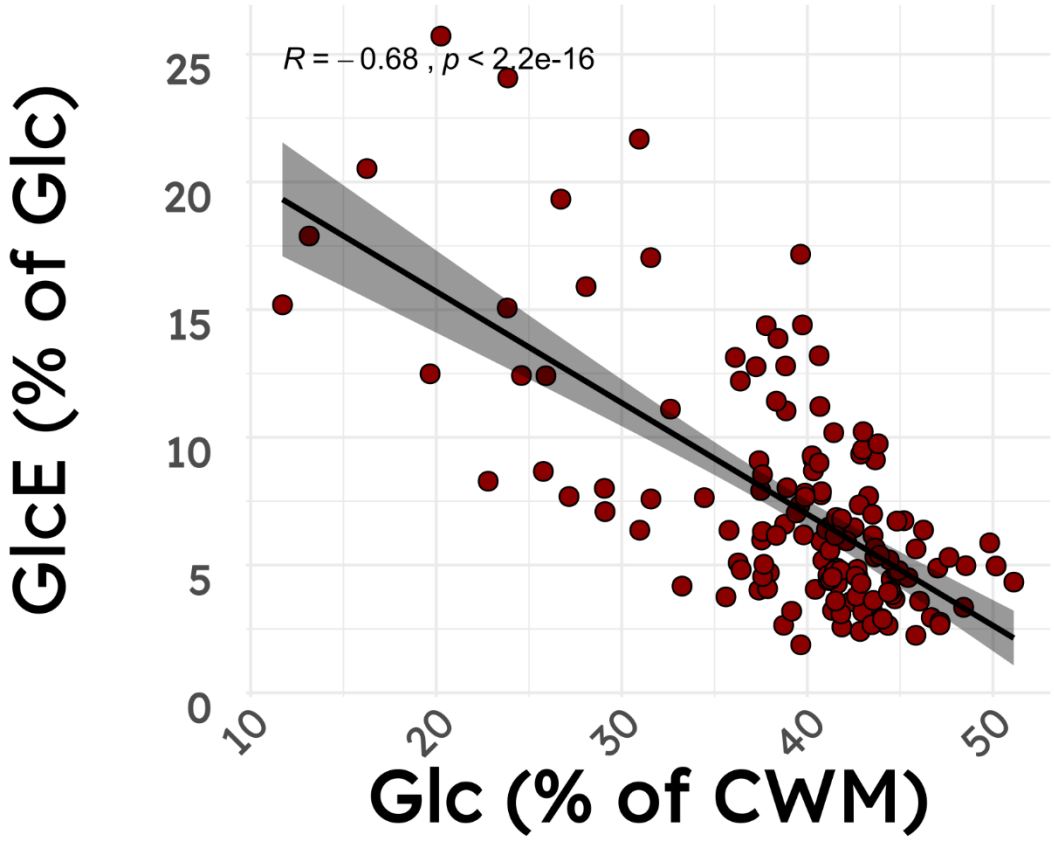


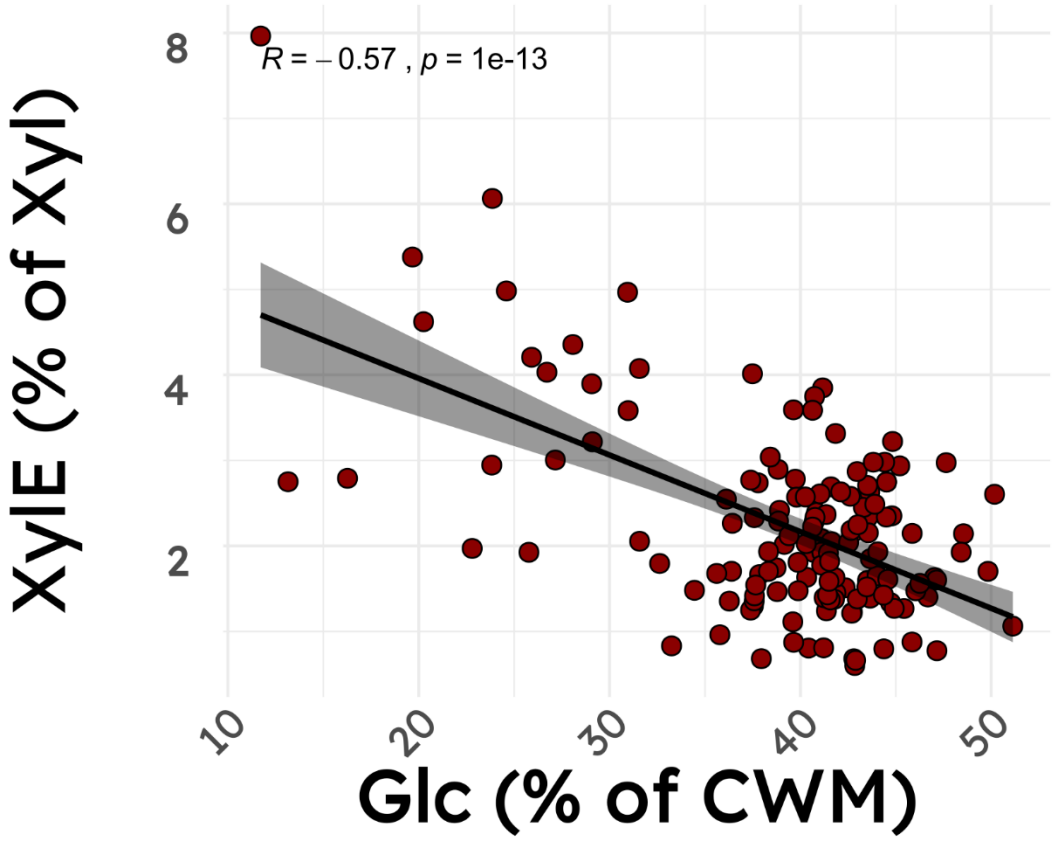
Appendix 4 Fig 1 Regression plot matrix of comparison, architecture and saccharification variables

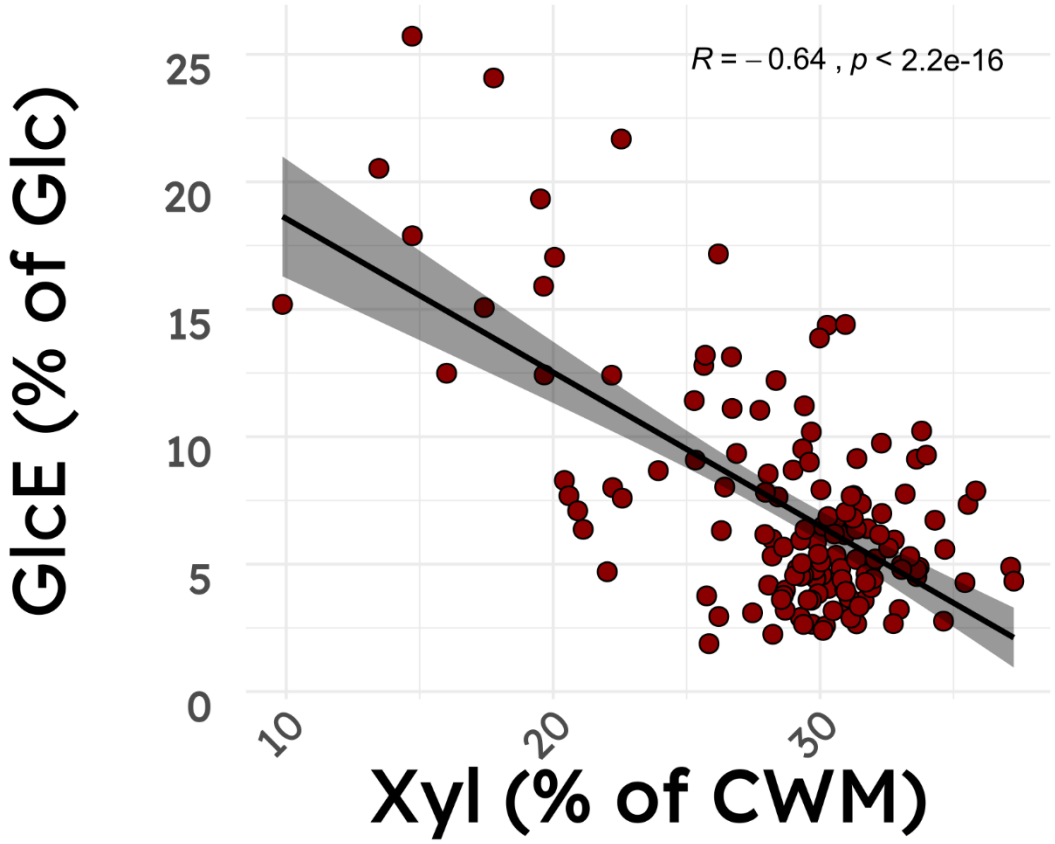

R = 0.57 , p = 2.3e-12
Gal (% CWM)
Ara (% CWM)
GlcE (%)
25
20
15
10
5
0
1
2
3
4
5

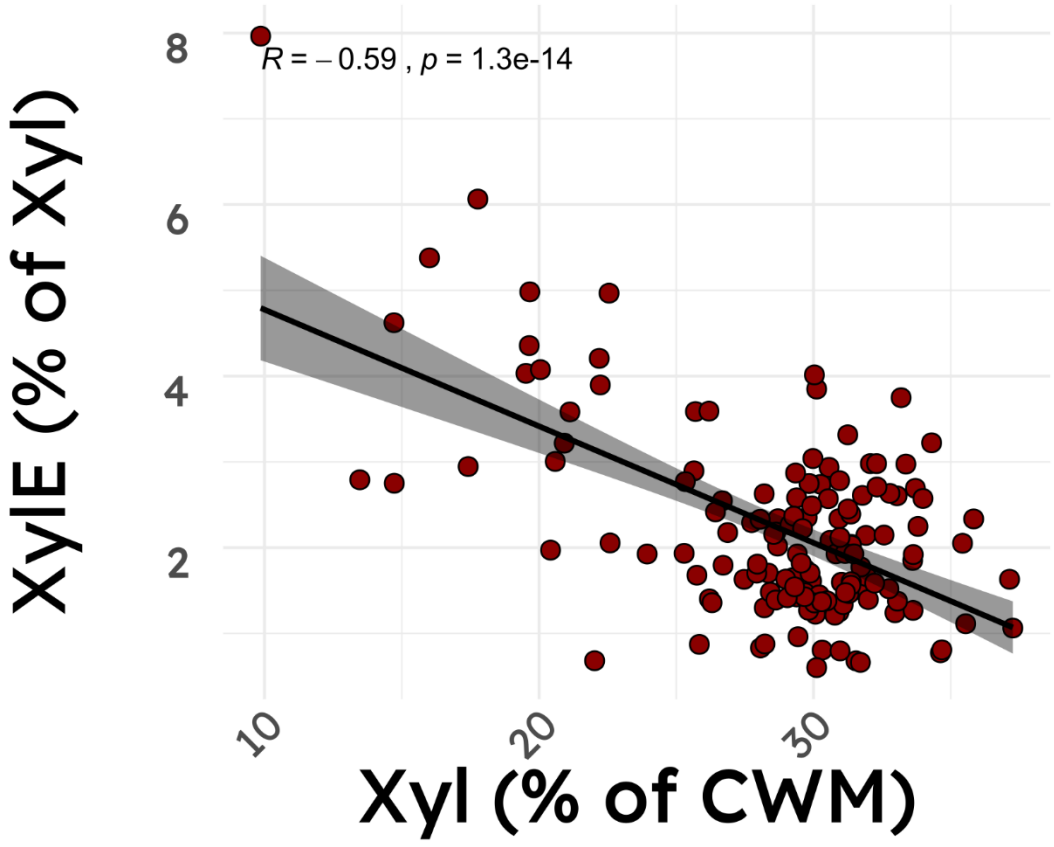

R = – 0.84 , p < 2.2e-16
Decoration (Xyl/Ara)
Ara (% of CWM)
GlcE (%)
25
20
15
10
5
0
10
20
1
2
3
4
5

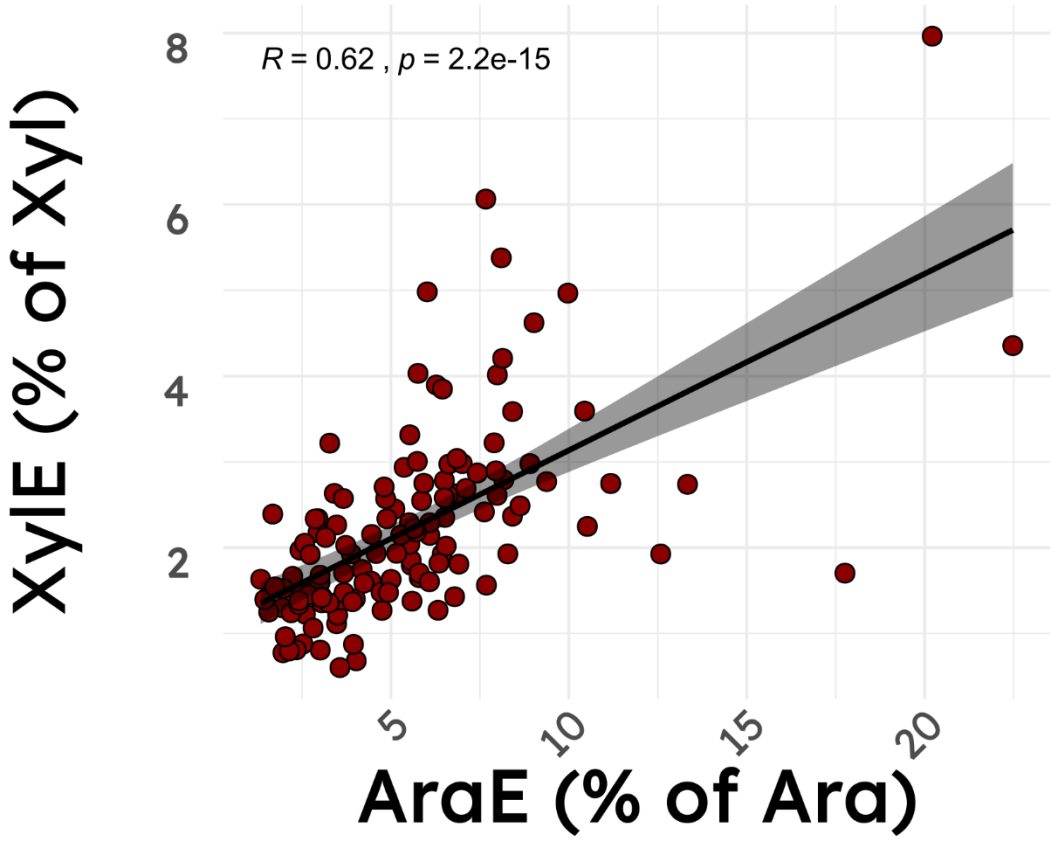

R = 0.87 , p < 2.2e-16
Xyl (% of CWM)
Glc (% of CWM)
GlcE (%)
25
20
15
10
5
10
20
30
40
50

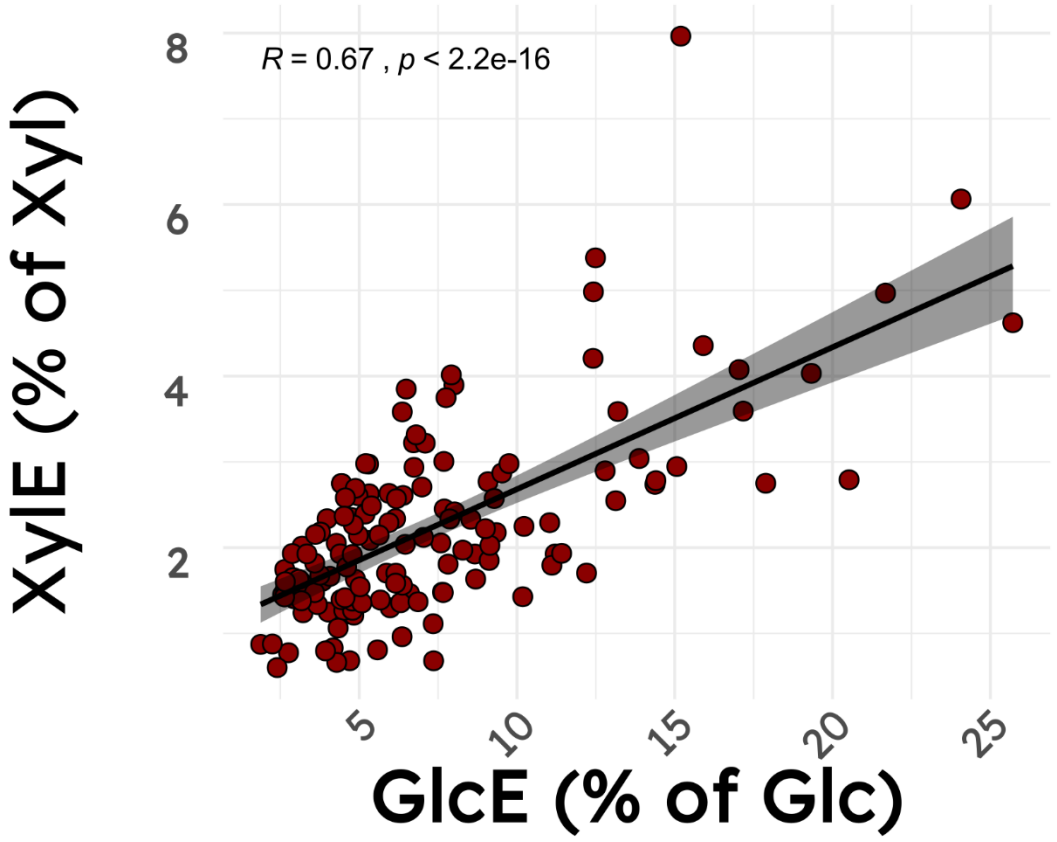

R = – 0.68 , p < 2.2e-16
GlcE (% of Glc)
Glc (% of CWM)
0
5
10
15
20
25
10
20
30
40
50

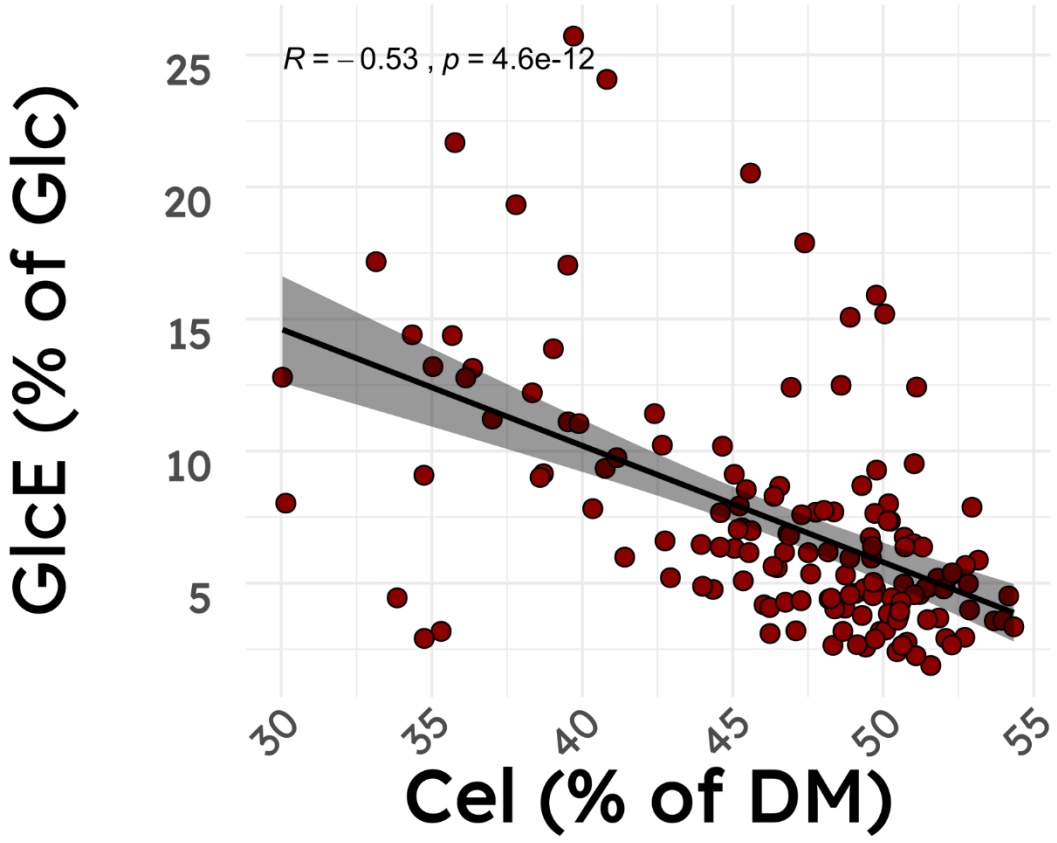

R = – 0.57 , p = 1e-13
XylE (% of Xyl)
Glc (% of CWM)
8
6
4
2
10
20
30
40
50

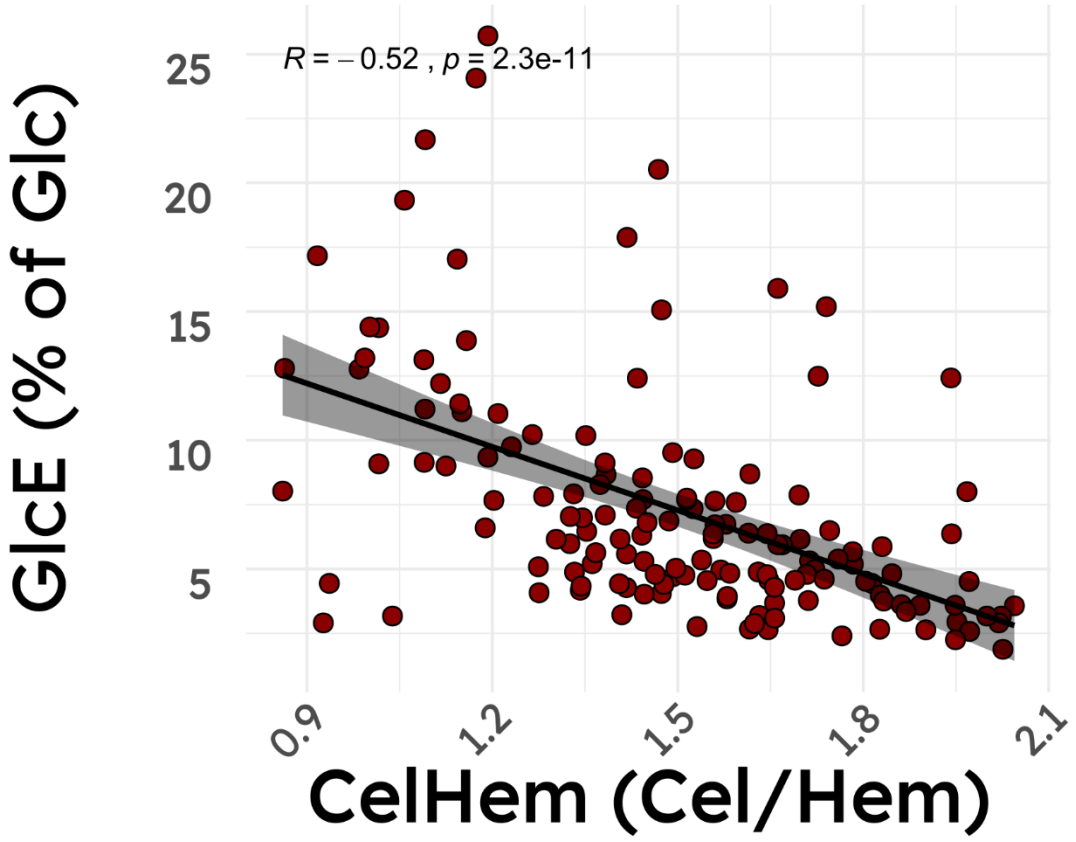

R = – 0.64 , p < 2.2e-16
GlcE (% of Glc)
Xyl (% of CWM)
25
20
15
10
5
0
10
20
30

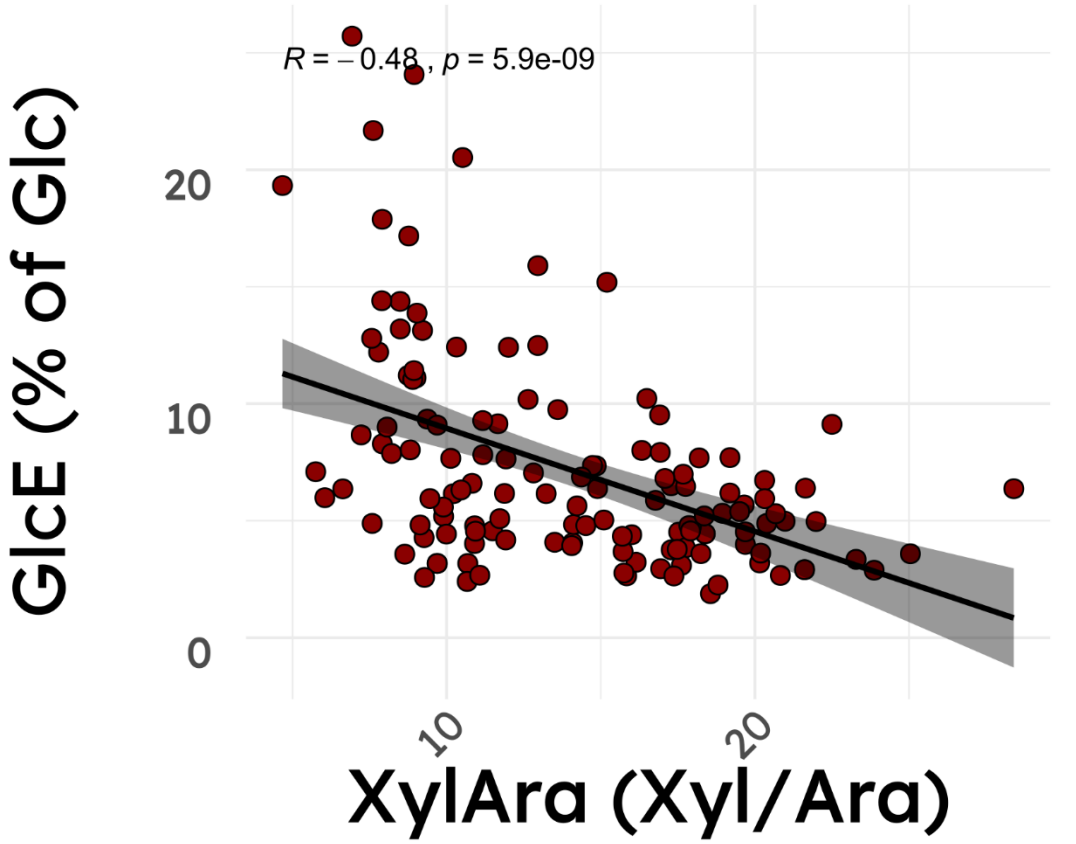

R = – 0.59 , p = 1.3e-14
XylE (% of Xyl)
8
6
4
2
10
20
30
Xyl (% of CWM)

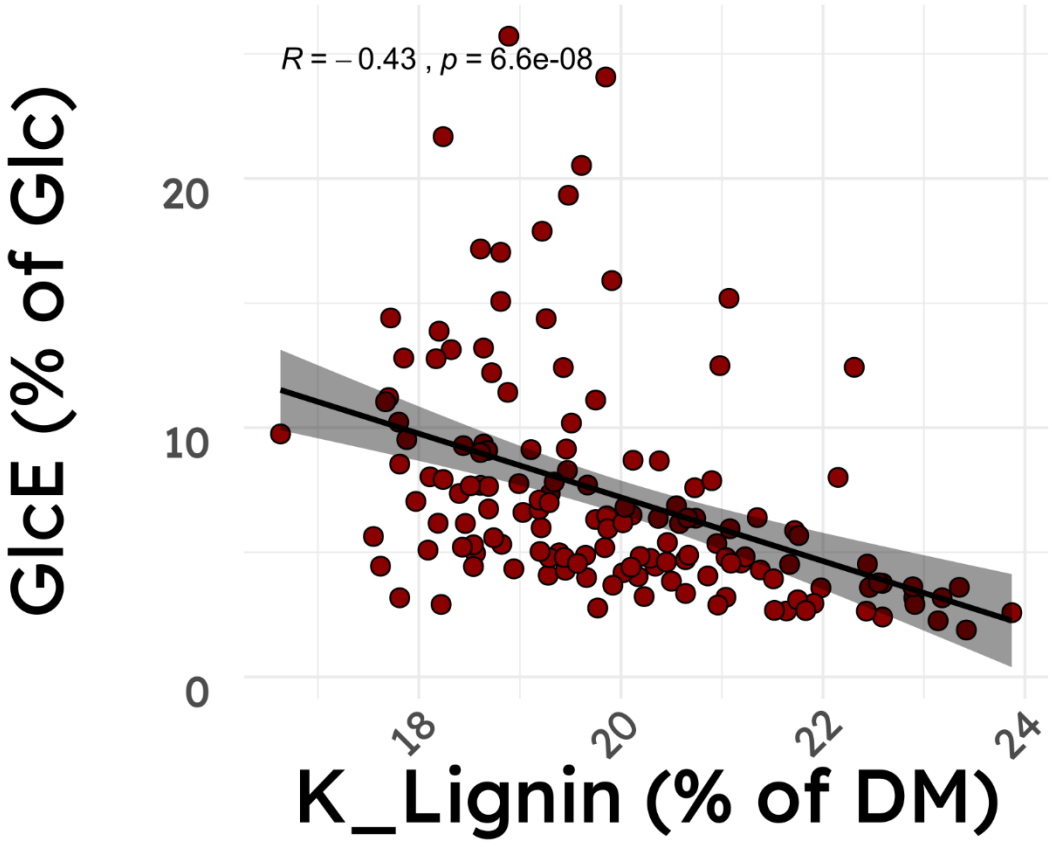

R = 0.62 , p = 2.2e-15
XylE (% of Xyl)
8
6
4
2
5
10
15
20
AraE (% of Ara)

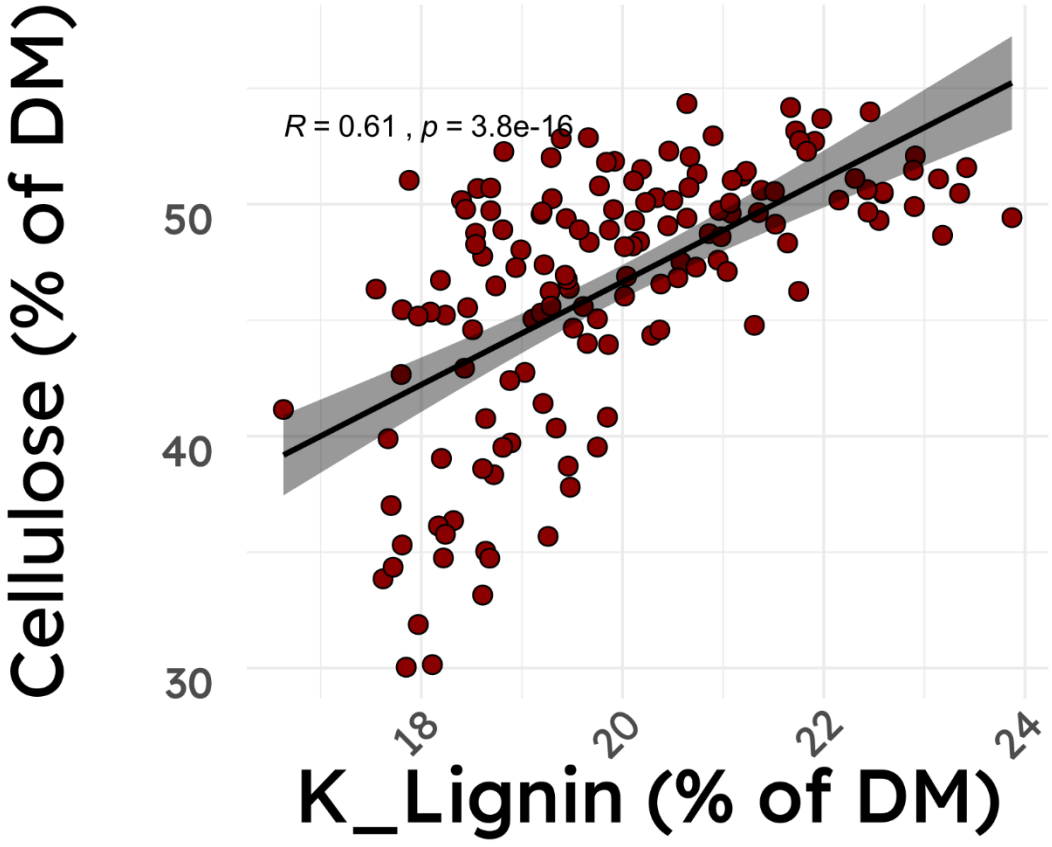

R = 0.67 , p < 2.2e-16
XylE (% of Xyl)
GlcE (% of Glc)
8
6
4
2
5
10
15
20
25

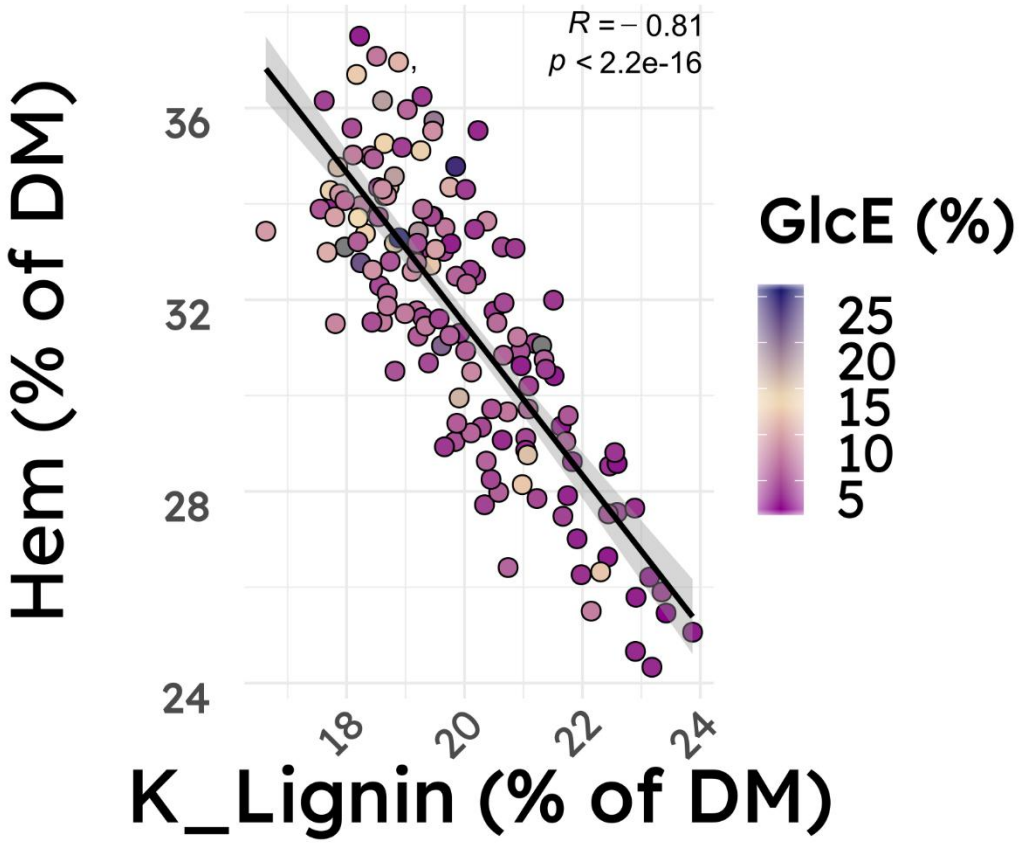

R = – 0.53 , p = 4.6e-12
GlcE (% of Glc)
Cel (% of DM)
25
20
15
10
5
30
35
40
45
50
55

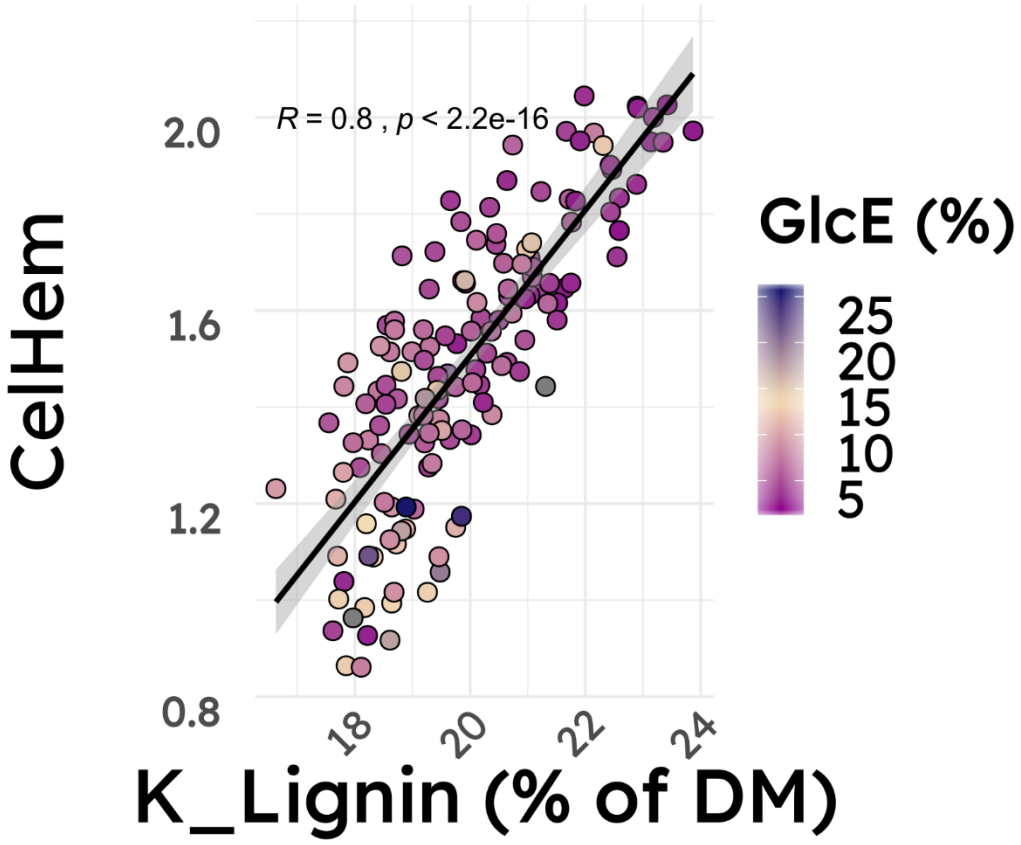

R = −0.52 , p = 2.3e-11
GlcE (% of Glc)
25
20
15
10
5
0.9
1.2
1.5
1.8
2.1
CelHem (Cel/Hem)

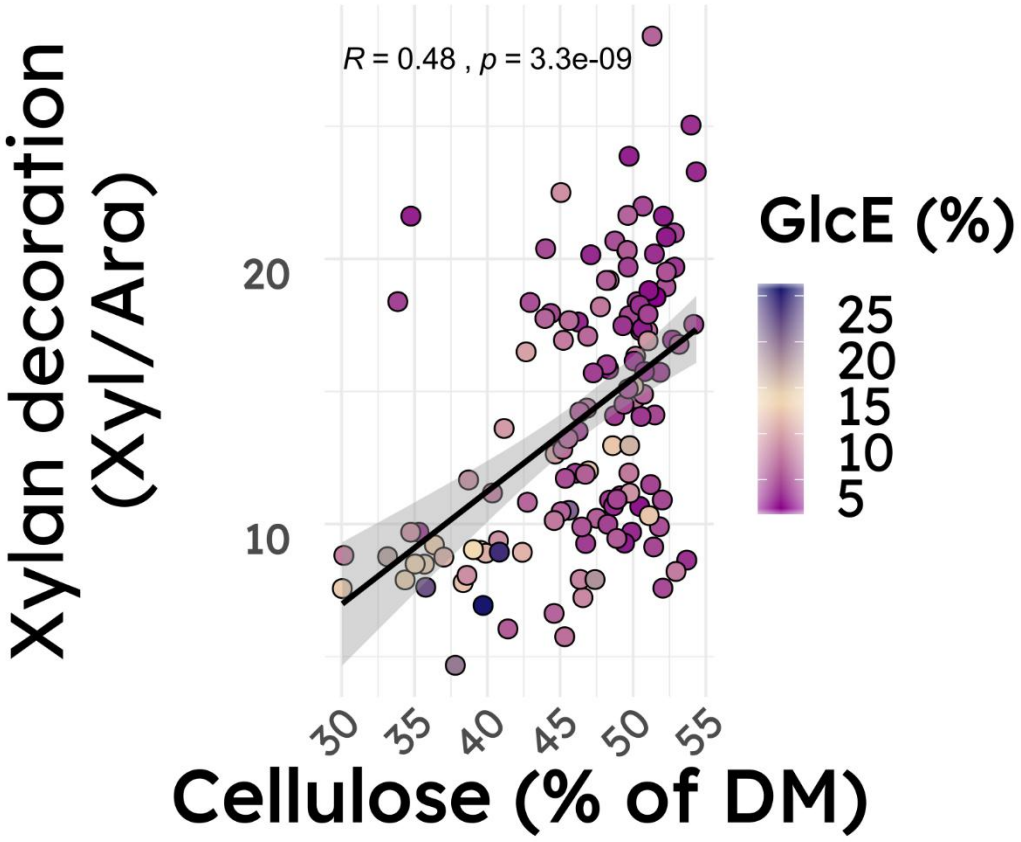

R = −0.48 , p = 5.9e-09
GlcE (% of Glc)
20
10
0
10
20
XylAra (Xyl/Ara)

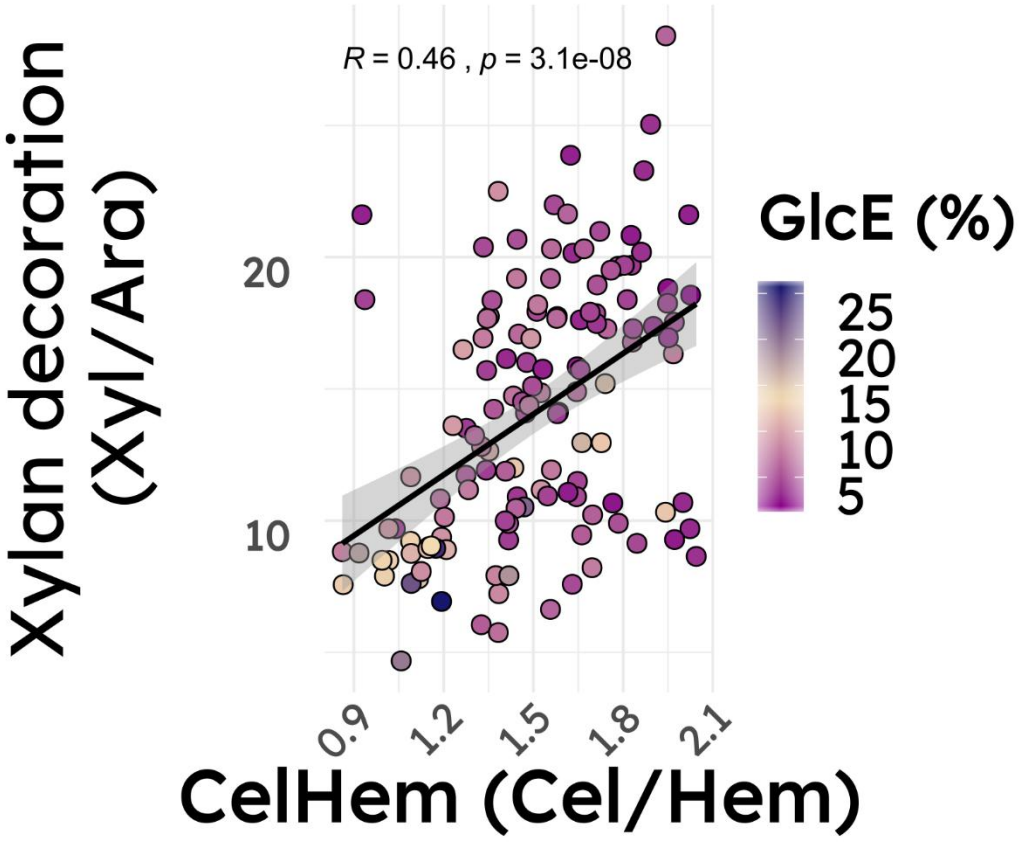

R = −0.43 , p = 6.6e-08
GlcE (% of Glc)
0
10
20
18
20
22
24
K_Lignin (% of DM)

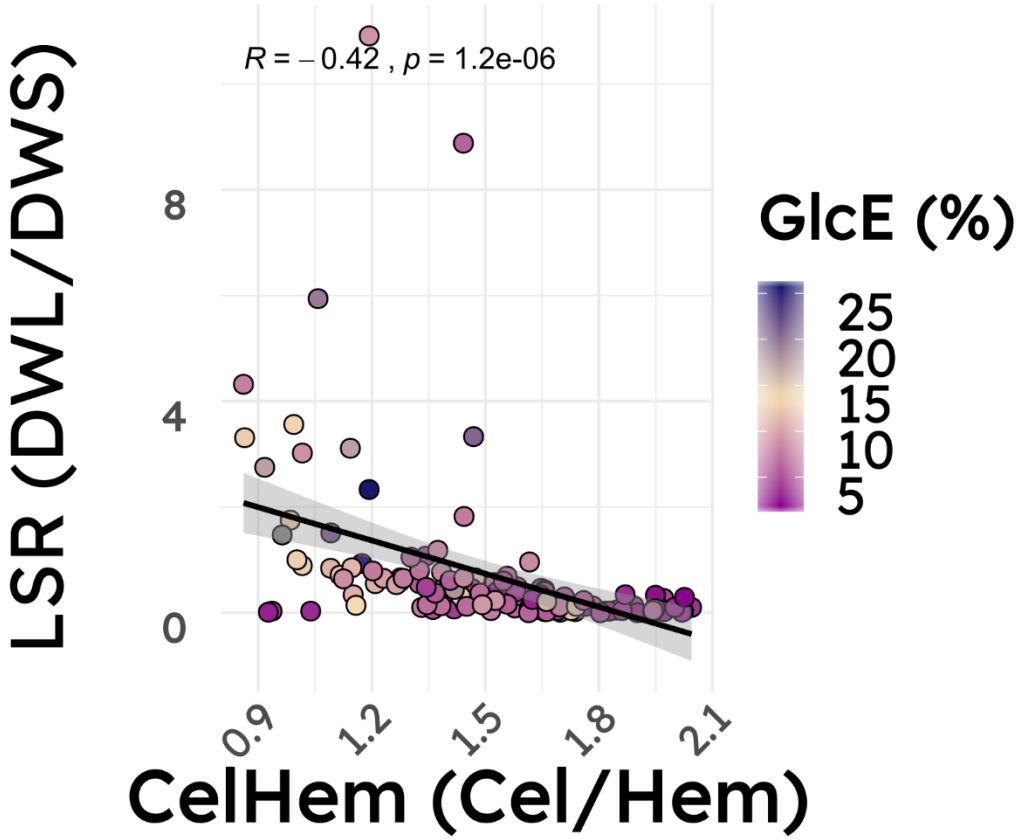

R = 0.61 , p = 3.8e-16
Cellulose (% of DM)
30
40
50
18
20
22
24
K_Lignin (% of DM)

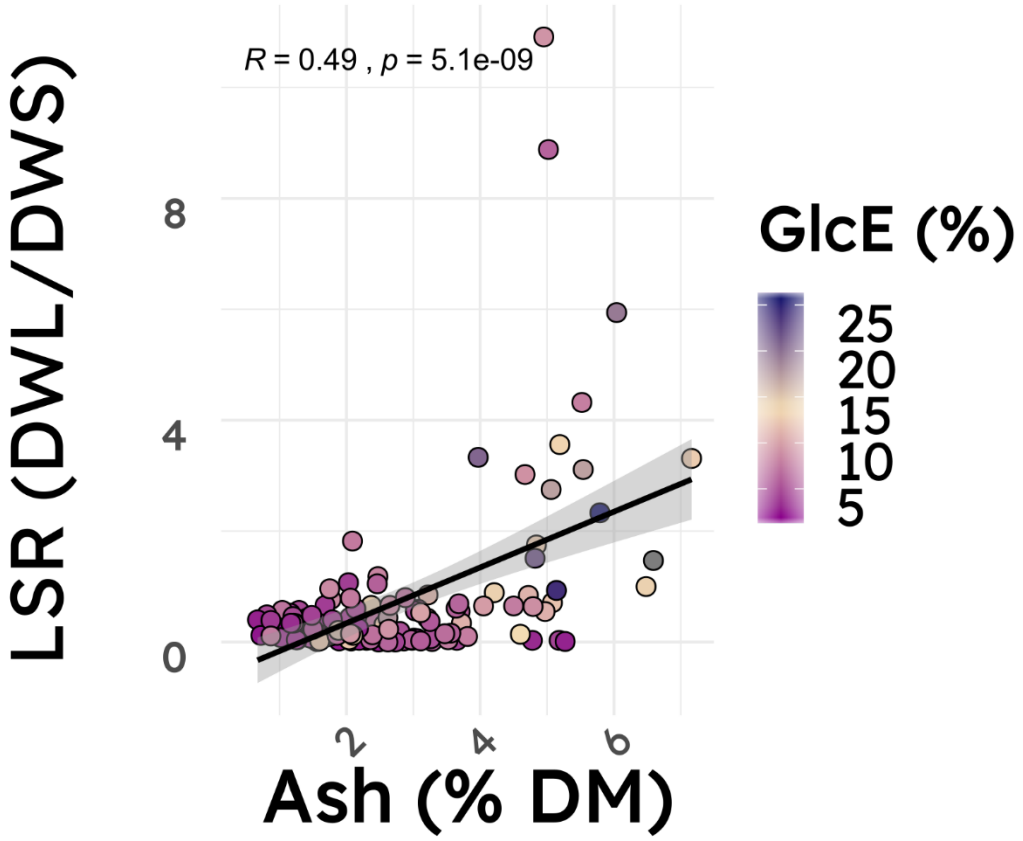

R = – 0.81
p < 2.2e-16
Hem (% of DM)
36
32
28
24
18
20
22
24
K_Lignin (% of DM)
GlcE (%)
25
20
15
10
5


R = 0.8 , p < 2.2e-16
CelHem
2.0
1.6
1.2
0.8
18
20
22
24
K_Lignin (% of DM)
GlcE (%)
25
20
15
10
5


R = 0.48 , p = 3.3e-09
Xylan decoration (Xyl/Ara)
20
10
30
35
40
45
50
55
Cellulose (% of DM)
GlcE (%)
25
20
15
10
5


R = 0.46 , p = 3.1e-08
Xylan decoration (Xyl/Ara)
20
10
0.9
1.2
1.5
1.8
2.1
CelHem (Cel/Hem)
GlcE (%)
25
20
15
10
5


R = – 0.42 , p = 1.2e-06
LSR (DWL/DWS)
8
4
0
0.9
1.2
1.5
1.8
2.1
CelHem (Cel/Hem)
GlcE (%)
25
20
15
10
5


R = 0.49 , p = 5.1e-09
LSR (DWL/DWS)
8
4
0
2
4
6
Ash (% DM)
GlcE (%)
25
20
15
10
5

Appendix E Chapter 4

E1 Summary statistics

Table E. 1 TGA data.

Variates associated with the thermogravimetric behaviour of the biomass in the 49 genotypes. Values are the average between three biological replicates (n = 3) and expressed as mean and standard deviation of the mean (sd)

Genetic group	Genotype	Onset temp.		Endset temp.		Step 1 mass loss		Step 2 mass loss		Step 3 mass loss	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Flor_Taiwan	Eufydd	293.18	4.36	360.91	0.69	-1.93	0.34	-29.49	2.17	-27.37	2.75
	Gawain	290.6	1.8	371.92	4.43	-1.99	0.35	-24.59	0.93	-31.78	2.56
	Geraint	289.52	5.71	356.18	3.51	-1.5	0.38	-28.94	0.77	-23.99	2.75
	Lancelot	291.14	4.39	376.75	0.86	-2.2	0.1	-24.67	4.05	-33.44	2.75
	Llyr	294.03	1.21	364.54	2.42	-2.34	0.12	-24.74	2.62	-31.33	3.78
	Percival	295.83	3.44	359.2	1.04	-1.5	0.22	-28	1.16	-26.56	1.5

Genetic group	Genotype	Onset temp.		Endset temp.		Step 1 mass loss		Step 2 mass loss		Step 3 mass loss	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Flor_Taiwan	Total	292.38	3.91	364.92	7.76	-1.91	0.4	-26.74	2.89	-29.08	4.14
Hyb.gig	Cilydd	305.1	13.44	377.71	9.48	-2.19	0.34	-25.11	5.04	-38.75	7.81
	Lamorak	308.09	1.31	380.63	1.46	-1.79	0.28	-24.6	0.73	-38.47	0.9
	Olwen	311.99	6.03	384.8	5.23	-2.14	0.48	-24.46	2.21	-40.24	3.46
Hyb.gig	Total	308.39	7.97	381.05	6.28	-2.04	0.38	-24.72	2.79	-39.15	4.37
Hyb.mx	Galahd	295.84	7.65	368.58	5.45	-2.02	0.23	-27.82	4.31	-33.03	4.5
	Tristan	289.38	7.46	370.94	2.59	-1.96	0.29	-28.15	3.29	-31.96	4.7
Hyb.mx	Total	292.61	7.63	369.76	4.03	-1.99	0.24	-27.98	3.44	-32.49	4.16
Hyb.sin	Cigfa	317.18	8.72	383.96	5.35	-2.31	0.34	-24.12	2.32	-42.7	4.32
	Hafgan	304.94	21.37	377.17	9.26	-2.21	0.27	-26.81	2.97	-37.03	8.28
	Nimue	313.91	12.91	384.54	7.11	-2.23	0.27	-23.9	3.37	-41.51	5.57
Hyb.sin	Total	312.01	14.32	381.89	7.33	-2.25	0.26	-24.94	2.89	-40.42	6.02
Lut	Blodeuwedd	301.47	8.58	364.38	10.48	-1.49	0.06	-26.8	5.3	-35.01	5.93

Genetic group	Genotype	Onset temp.		Endset temp.		Step 1 mass loss		Step 2 mass loss		Step 3 mass loss	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
	Bran	296.98	9.23	363.6	4.69	-1.5	0.06	-27.91	7.15	-33.2	6.6
	Eigyr	292.84	3.65	366.66	4.8	-1.52	0.11	-28.43	0.94	-33.01	1.01
	Goronwy	296.2	9.75	356.18	10.17	-1.43	0.06	-29.39	7.57	-31.3	7.92
	Govannon	300.75	10.17	364.34	4.68	-1.35	0.1	-28.29	6.23	-34.24	5.34
	Gwalchmei	298.82	6.59	364.99	9.29	-1.47	0.12	-27.7	8.23	-34.94	7.09
	Gwydion	303.31	5.39	374.57	2.35	-1.33	0.1	-23.31	4.54	-38.72	4.18
	Kay	296.09	7.24	367.39	3.08	-1.39	0.1	-28.32	4.18	-32.1	4.45
	Lleu	295.83	6.68	366.68	4.19	-1.79	0.16	-28.09	1.41	-33.25	1.85
	Mabon	305.89	10.68	371.44	10.57	-1.45	0.14	-26.71	6.13	-37.3	7.15
	Math	297.75	5.05	364.21	8.71	-1.43	0.04	-28.06	7.17	-33.61	6.04
	Owain	302.75	19.75	373.09	11.93	-2.34	0.28	-28.17	6.23	-35.7	8.98
Lut	Total	298.94	8.69	366.23	8.01	-1.55	0.29	-27.72	5.03	-34.24	5.35
Sacc_rob	Aranrhod	302.49	6.95	376.8	5.59	-1.65	0.83	-24.75	4.98	-39.49	6.44

Genetic group	Genotype	Onset temp.		Endset temp.		Step 1 mass loss		Step 2 mass loss		Step 3 mass loss	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
	Beli	302.79	9.08	372.93	1.31	-2.23	0.22	-23.95	4.48	-36.72	3.82
	Culhwch	312.04	12.4	381.32	5.26	-1.48	0.12	-24.35	2.11	-41.58	3.67
	Gaheris	323.29	4.11	386	5.32	-2.28	0.41	-21.85	1.33	-45.98	2.87
	Gwenhwyfar	301.84	12.82	374.61	14.91	-2.36	0.43	-26.44	7.57	-37.78	10.03
	Lunete	316.78	2.35	381.78	3.64	-2.06	0.2	-21.12	3.52	-43.46	1.43
	Modred	306.34	8.27	380.67	4.87	-2.01	0.24	-21.89	2.61	-40.92	3.13
	Nudd	296.53	8.13	377.85	4.64	-2	0.14	-24.87	2.66	-35.57	3.93
	Pryderi	299.44	14.24	368.85	10.1	-1.96	0.29	-26.4	4.99	-33.98	7.56
	Pwyll	307.32	8.87	379.31	3.48	-1.91	0.25	-24.38	1.07	-38.83	2.55
	Rhiannon	299.64	9.27	379.42	9.17	-2.29	0.38	-26.4	4.42	-37.14	6.54
Sacc_rob	Total	306.23	11.1	378.14	7.51	-2.02	0.41	-24.22	3.83	-39.22	5.6
Sin (EMI/PRI)	Luned	302.32	11.83	378.25	7.05	-1.84	0.39	-27.47	4.32	-35.39	3.32

Genetic group	Genotype	Onset temp.		Endset temp.		Step 1 mass loss		Step 2 mass loss		Step 3 mass loss	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Sin (EMI/PRI)	Total	302.32	11.83	378.25	7.05	-1.84	0.39	-27.47	4.32	-35.39	3.32
Sin_N_Japan	Aeron	284.21	3.19	370.1	6.48	-1.67	0.17	-28.28	2.65	-28.73	3.44
	Dylan	307.51	4.58	381.93	3.33	-2.1	0.27	-25.41	1.82	-36.68	1.68
	Enid	311.94	8.4	383.35	0.33	-2.11	0.19	-24.12	0.29	-40.33	2.73
	Gareth	284.83	8.9	361.39	13.11	-2.59	1.32	-30.95	4.96	-25.58	5.53
	Peredur	306.98	8.16	379.58	3.78	-2	0.28	-25.48	2.45	-36.42	2.02
Sin_N_Japan	Total	299.09	13.8	375.27	10.43	-2.09	0.61	-26.85	3.51	-33.55	6.36
Sin_S_Japan	Angharad	287.17	20.16	379.07	6.16	-1.86	0.36	-26.18	2.83	-31.61	5.53
	Bedivere	315.04	5.75	382.79	3.57	-2.33	0.32	-25.27	1.08	-40.59	2.73
	Urien	291.38	8.59	367.39	2.59	-2.13	0.34	-28.11	3.46	-30.71	5.06
Sin_S_Japan	Total	297.86	17.25	376.42	7.92	-2.11	0.36	-26.52	2.62	-34.3	6.19
Sin_Taiwan	Bors	313.23	4.03	382.81	2.02	-1.94	0.25	-24.85	1.12	-39.17	1.5

Genetic group	Genotype	Onset temp.		Endset temp.		Step 1 mass loss		Step 2 mass loss		Step 3 mass loss	
		mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
	Uther	295.79	7.43	380.34	2.22	-1.99	0.14	-26.04	1.19	-35.65	1.71
	Yspaddaden	305.8	11.17	381.53	6.59	-2.17	0.4	-26.85	3.05	-38.08	5.47
Sin_Taiwan	Total	304.94	10.32	381.56	3.78	-2.03	0.27	-25.91	1.94	-37.63	3.35

E2 ANOVA tables

Table E. 2 Test for the effect of genotype and genetic group on TG parameters

The effect of genotype and genetic group (independently) was tested on the parameters of the tg curve. When the p value was lower than 0.05 a Tukey HSD posthoc test was used to detect the difference.

Variable	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power
Onset temp.	Genotype	48	11566.7	240.973	2.753	0	0.577	0.577	0.366	0.366	0.367	1.167	1
	Residuals	97	8490.613	87.532									
	Genetic group	9	4667.521	518.613	4.583	0	0.233	0.233	0.181	0.181	0.182	0.551	0.999
	Residuals	136	15389.79	113.16									

Endset temp.	Genotype	48	9857.679	205.368	4.85	0	0.706	0.706	0.559	0.559	0.56	1.549	1
	Residuals	97	4107.761	42.348									
	Genetic group	9	5894.121	654.902	11.04	0	0.422	0.422	0.382	0.382	0.384	0.855	1
	Residuals	136	8071.32	59.348									
Mass loss Step 1	Genotype	48	15.775	0.329	2.808	0	0.582	0.582	0.373	0.373	0.374	1.179	1
	Residuals	97	11.353	0.117									
	Genetic group	9	7.215	0.802	5.475	0	0.266	0.266	0.216	0.216	0.217	0.602	1
	Residuals	136	19.913	0.146									
Mass loss Step 2	Genotype	48	653.905	13.623	0.835	0.753	0.292	0.292	-0.057	-0.057	-0.058	0.643	0.966
	Residuals	97	1582.184	16.311									

	Genetic group	9	281.521	31.28	2.176	0.027	0.126	0.126	0.068	0.068	0.068	0.38	0.899
	Residuals	136	1954.569	14.372									
Mass loss													
Step 3	Genotype	48	3176.388	66.175	2.72	0	0.574	0.574	0.361	0.361	0.363	1.16	1
	Residuals	97	2360.312	24.333									
	Genetic group	9	1758.398	195.378	7.033	0	0.318	0.318	0.271	0.271	0.272	0.682	1
	Residuals	136	3778.302	27.782									

Figures

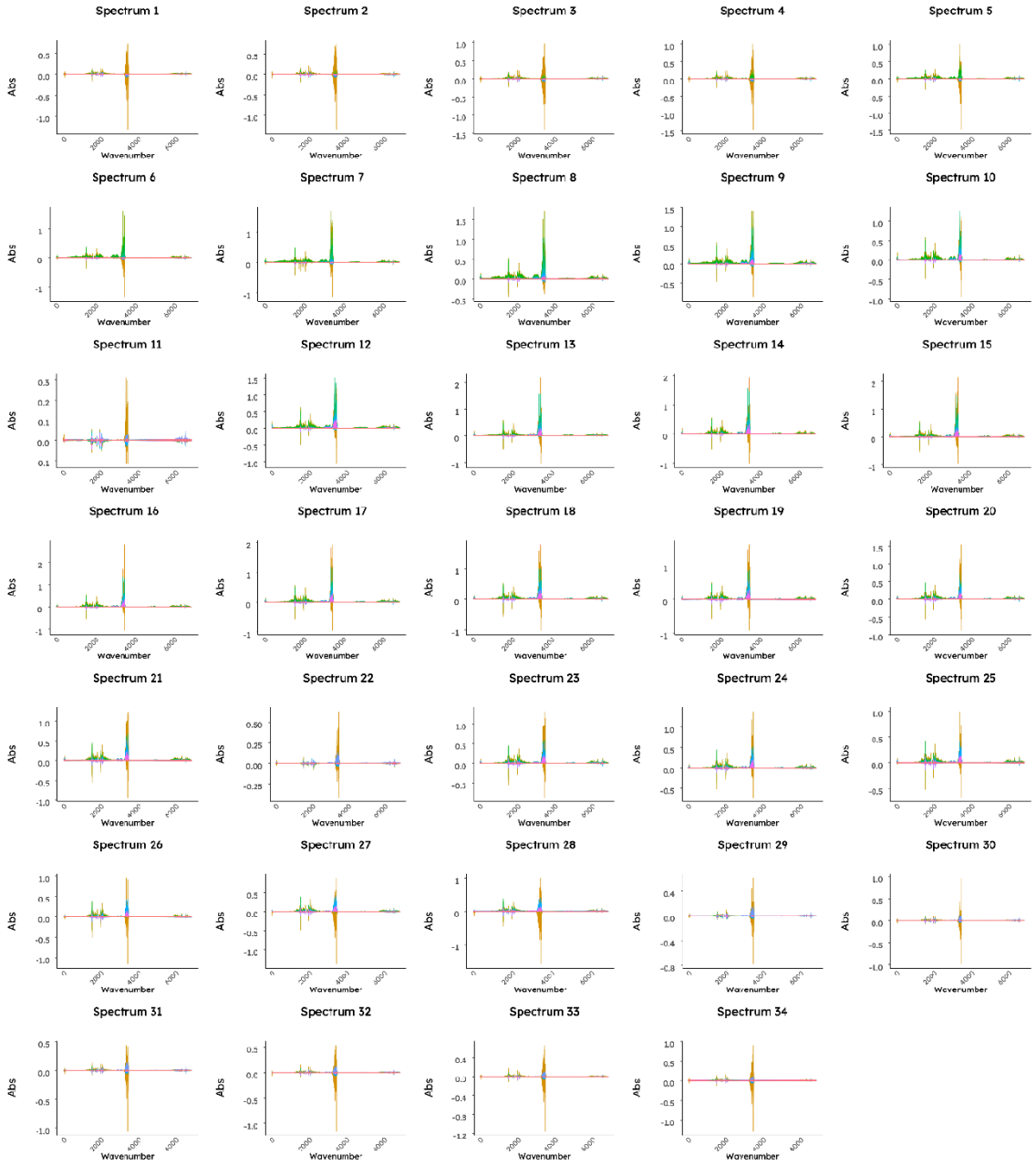


Figure E. 1 Check of FTIR spectra shift

The shift in the FTIR spectra was visually evaluated. The spectra from all the samples for each time point were overlapped to detect differences related to peak-shift.

Appendix F Chapter 5

F1 Summary statistics

Table F. 1 Summary of the amount of monosaccharides enzymatically released

The amount of glucose (GlcE) and xylose (XylE) enzymatically released expressed as mean an standard deviation of the mean (mean±sd) between three biological replicates (n = 3). The analysis was performed on 4 technical replicates.

Genotype	Genetic group	Plant portion	GlcE		XylE	
			mean±sd	median	mean±sd	median
Gawain	Flor_Taiwan	Leaf	13.91±3.65	15.64	4.81±0.2	4.73
		Stem	10.27±2.26	9.14	6.78±0.74	6.46
Geraint	Flor_Taiwan	Leaf	14.28±9.14	11.28	5.44±1.7	5.11
		Stem	6.84±0.66	7.04	5.77±1.42	6.04
Lancelot	Flor_Taiwan	Leaf	17.47±2.04	18.59	6.09±0.53	6.02
		Stem	6.83±1.66	6.83	5.66±0.81	5.66
Percival	Flor_Taiwan	Leaf	13.49±9.93	13.58	6.99±1.48	7
		Stem	5.41±4.52	7.52	6.63±1.39	7.13
Lamorak	Hyb.gig	Leaf	19.69±1.83	19.26	6.65±0.6	6.5
		Stem	5.48±0.84	5.79	4.46±0.24	4.45
Galahad	Hyb.mx	Leaf	17.06±3.01	15.41	5.55±0.68	5.77
		Stem	10.56±1.77	10.74	5.08±0.85	5.39

Genotype	Genetic group	Plant portion	GlcE		XylE	
			mean±sd	median	mean±sd	median
Tristan	Hyb.mx	Leaf	24.64±1.49	24.95	6.91±0.49	6.78
		Stem	6.62±3.06	5.2	4.72±1.34	4.7
Kay	Lut	Leaf	23.79±1.53	22.99	7.33±0.39	7.13
		Stem	8.46±2.09	7.41	5.49±0.35	5.57
Gaheris	Sacc_rob	Leaf	17.59±3.27	18.4	5.6±0.43	5.74
		Stem	5.08±0.87	4.82	3.49±0.2	3.54
Gareth	Sin_N_Japan	Leaf	18.29±1.36	17.73	5.3±0.51	5.33
		Stem	5.5±1.48	5.41	4.28±0.38	4.12
Bedivere	Sin_S_Japan	Leaf	18.72±3.54	20.51	5.38±1.13	4.8
		Stem	5.32±1.66	6.03	4.42±0.18	4.39
Bors_TY	Sin_Taiwan	Leaf	12.34±3.11	11.31	4.56±0.32	4.52
		Stem	7.53±1.1	7.12	5.29±0.51	5.14

Table F. 2 Summary of matrix monosaccharide content – Part 1

Hemicellulose was extraxted using tri-fluoro-acetic acid as described in subsection 5.2.5.4. Values are monosaccharide content in the supernatant espressed as a percentage of cell wall material. Mean and sd are the average and the standard deviation between the three biological replicates.

		Plant	Ara		Fuc		Gal	
Genotype	Genetic group	portion	mean±sd	median	mean±sd	median	mean±sd	median
Gawain	Flor_Taiwan	Leaf	3.4±0.36	3.52	0.21±0.18	0.16	1.22±0.04	1.22
		Stem	2.27±0.1	2.28	1.08±1.51	0.26	0.84±0.1	0.87
Geraint	Flor_Taiwan	Leaf	4.14±0.76	4.34	0.16±0.2	0.05	1.37±0.07	1.39
		Stem	2.63±0.47	2.81	0.57±0.88	0.07	0.84±0.26	0.93
Lancelot	Flor_Taiwan	Leaf	3.59±0.76	3.95	0.5±0.67	0.18	1.31±0.28	1.25
		Stem	2.19±0.13	2.19	0.23±0.15	0.23	0.6±0.08	0.6
Percival	Flor_Taiwan	Leaf	3.52±0.9	3.05	0.43±0.59	0.12	1.11±0.21	1.1
		Stem	2.39±0.47	2.25	0.97±1.02	0.76	0.87±0.21	0.8

Lamorak	Hyb.gig	Leaf	3.78±0.07	3.77	0.55±0.81	0.11	1.25±0.1	1.23
		Stem	2.22±0.33	2.31	0.06a□0.02	0.06	0.78±0.18	0.8
Galahad	Hyb.mx	Leaf	3.18±0.88	2.91	0.26±0.16	0.23	1.02±0.09	1.02
		Stem	1.98±0.15	1.95	0.14±0.09	0.13	0.58±0.08	0.55
Tristan	Hyb.mx	Leaf	3.34±0.76	3.06	0.16±0.12	0.11	1.13±0.1	1.12
		Stem	2.48±0.81	2.46	0.41±0.61	0.06	0.78±0.32	0.77
Kay	Lut	Leaf	3.01±0.83	3.23	2.05±1.65	2.78	1.08±0.42	1.02
		Stem	2.45±0.33	2.39	0.72±0.98	0.72	0.66±0.13	0.69
Gaheris	Sacc_rob	Leaf	2.75±0.18	2.74	0.2±0.21	0.1	0.82±0.04	0.8
		Stem	1.79±0.25	1.67	0.08±0.05	0.09	0.54±0.16	0.52
Gareth	Sin_N_Japan	Leaf	3.87±0.91	3.61	0.06±0.01	0.06	1.24±0.28	1.3
		Stem	2.23±0.39	2.09	0.09±0.06	0.1	0.63±0.12	0.58
Bedivere	Sin_S_Japan	Leaf	2.47±0.43	2.57	1.07±0.85	1.48	0.88±0.21	0.96
		Stem	1.86±0.05	1.86	0.26±0.35	0.1	0.63±0.14	0.71
Bors_TY	Sin_Taiwan	Leaf	2.9±0.28	2.89	0.32±0.23	0.45	0.85±0.06	0.85

		Stem	2.52±0.34	2.45	0.3±0.37	0.14	0.71±0.13	0.74

Table F. 3 Summary of matrix monosaccharide content – Part 2

Hemicellulose was extraxted using tri-fluoro-acetic acid as described in subsection 5.2.5.4. Values are monosaccharide content in the supernatant espressed as a percentage of cell wall material. Mean and sd are the average and the standard deviation between the three biological replicates.

		Plant	Man		Rha		Xyl	
Genotype	Genetic group	portion	mean±sd	median	mean±sd	median	mean±sd	median
Gawain	Flor_Taiwan	Leaf	3.72±0.31	3.77	0.26±0.04	0.28	23.3±0.45	23.55
		Stem	4.7±1.3	4.48	0.19±0.05	0.17	25.49±1.09	25.32
Geraint	Flor_Taiwan	Leaf	4.54±0.61	4.61	0.35±0.08	0.35	24.85±0.76	24.52
		Stem	4.49±1.21	5.18	0.31±0.22	0.34	26.34±1.08	25.76
Lancelot	Flor_Taiwan	Leaf	3.98±1.12	3.35	0.39±0.2	0.3	24.71±0.93	24.94

		Stem	3.6±0.66	3.6	0.1±0.02	0.1	26±0.61	26
Percival	Flor_Taiwan	Leaf	3.67±0.83	3.35	0.17±0.02	0.17	24.42±3.23	23.33
		Stem	4.99±1.88	5.09	0.18±0.23	0.05	24.79±1.32	24.33
Lamorak	Hyb.gig	Leaf	3.22±0.75	3.41	0.23±0.09	0.23	25.86±1.34	25.27
		Stem	3.93±0.86	3.7	0.26±0.17	0.33	23.98±1.13	24.19
Galahad	Hyb.mx	Leaf	2.99±0.56	2.87	0.16±0.03	0.17	26.16±1.55	26.26
		Stem	2.9±0.6	2.61	0.11±0.09	0.11	26.41±2.38	25.52
Tristan	Hyb.mx	Leaf	3.83±0.85	4.08	0.22±0.12	0.17	26.81±1.24	27.01
		Stem	4.03±0.9	3.68	0.44±0.35	0.52	27.57±1.85	27.01
Kay	Lut	Leaf	5.42±2.03	6.11	0.39±0.25	0.41	26.24±1.33	26.77
		Stem	3.85±0.81	3.92	0.27±0.13	0.31	27.13±1.02	26.75
Gaheris	Sacc_rob	Leaf	3.63±0.97	3.56	0.12±0.03	0.13	26.39±0.43	26.21
		Stem	3.04±0.85	3.1	NA±NA	NA	25.88±2.31	25.25
Gareth	Sin_N_Japan	Leaf	4.56±2.77	2.99	0.3±0.29	0.16	25.06±1.28	25.27
		Stem	3.88±0.99	4.3	0.17±0.11	0.15	25.21±1.91	26.1

Bedivere	Sin_S_Japan	Leaf	3.25±1.09	3.64	0.32±0.12	0.33	20.82±2.74	20.13
		Stem	3.55±0.34	3.46	0.14±0.12	0.11	26.23±2.09	26.29
Bors_TY	Sin_Taiwan	Leaf	3.51±0.55	3.47	0.21±0.05	0.22	26.91±0.66	26.71
		Stem	3.63±1.02	3.66	0.25±0.26	0.14	24.57±1.89	24.52

Table F. 4 Cellulose and hemicellulose content in the cell wall.

Genotype	Genetic group	Plant portion	Cellulose		Hemicellulose	
			mean±sd	median	mean±sd	median
Gawain	Flor_Taiwan	Leaf	36.36±1.51	36.63	16.04±0.56	16.11
		Stem	41.23±1.32	41.58	16.35±0.44	16.44
Geraint	Flor_Taiwan	Leaf	37.57±1.23	38.22	18.05±1.29	17.47
		Stem	41.25±1.14	40.62	17.32±2.27	17.97
Lancelot	Flor_Taiwan	Leaf	35.52±1.36	35.18	17.99±0.64	17.65
		Stem	39.5±1.19	39.5	12.85±11.47	16.49
Percival	Flor_Taiwan	Leaf	37.89±5.17	36.24	17.75±1.74	18.69
		Stem	42.25±0.87	42.53	17.76±1.14	17.26
Lamorak	Hyb.gig	Leaf	38.28±1.5	38.23	15.93±3.41	17.62
		Stem	42.5±1.09	42.03	17.32±0.71	17.35
Galahad	Hyb.mx	Leaf	37.07±0.21	37.19	17.25±1.14	17.14
		Stem	40.1±3.08	40.22	15.71±2.47	16.04
Tristan	Hyb.mx	Leaf	36.45±2.46	37.39	18.46±1.65	17.69
		Stem	40.8±1.03	41.37	18.53±0.73	18.46
Kay	Lut	Leaf	40.36±1.77	40.45	15.61±1.06	16.16
		Stem	41.1±1.1	41.68	17.64±1.82	18.58
Gaheris	Sacc_rob	Leaf	40.12±1.2	40.25	16.03±0.98	16.32
		Stem	42.35±1.59	42.75	16.37±0.47	16.43

Gareth	Sin_N_Japan	Leaf	37±1.18	37.33	17.66±1.09	18.01
		Stem	40.88±1.47	40.2	17.52±0.53	17.22
Bedivere	Sin_S_Japan	Leaf	31.65±3.99	29.53	14.97±3.13	15.65
		Stem	39.78±2.18	40.54	17.57±0.22	17.61
Bors_TY	Sin_Taiwan	Leaf	36.14±1.68	35.28	17.75±2.34	17.35
		Stem	38.91±0.98	38.66	19.47±2.8	18.86

Table F. 5 Cell wall architecture.

Genotype	Genetic group	Plant portion	Ara/Xyl		Cel/Hem	
			mean±sd	median	mean±sd	median
Gawain	Flor_Taiwan	Leaf	0.15±0.02	0.15	2.27±0.07	2.25
		Stem	0.09±0.01	0.09	2.52±0.01	2.53
Geraint	Flor_Taiwan	Leaf	0.17±0.03	0.18	2.09±0.21	2.19
		Stem	0.1±0.02	0.1	2.41±0.32	2.37
Lancelot	Flor_Taiwan	Leaf	0.15±0.03	0.16	1.98±0.13	1.99
		Stem	0.08±0	0.08	2.1±0.49	2.1
Percival	Flor_Taiwan	Leaf	0.14±0.02	0.14	2.14±0.3	2.3
		Stem	0.1±0.01	0.09	2.39±0.19	2.49
Lamorak	Hyb.gig	Leaf	0.15±0.01	0.15	2.5±0.7	2.1
		Stem	0.09±0.01	0.1	2.46±0.05	2.43

Galahad	Hyb.mx	Leaf	0.12±0.03	0.11	2.15±0.13	2.17
		Stem	0.07±0	0.07	2.59±0.42	2.4
Tristan	Hyb.mx	Leaf	0.12±0.03	0.12	1.98±0.15	1.9
		Stem	0.09±0.02	0.09	2.21±0.14	2.24
Kay	Lut	Leaf	0.11±0.03	0.12	2.59±0.1	2.6
		Stem	0.09±0.01	0.08	2.35±0.29	2.22
Gaheris	Sacc_rob	Leaf	0.1±0.01	0.11	2.51±0.17	2.45
		Stem	0.07±0	0.07	2.59±0.15	2.54
Gareth	Sin_N_Japan	Leaf	0.15±0.03	0.14	2.1±0.17	2.11
		Stem	0.09±0.01	0.08	2.33±0.02	2.33
Bedivere	Sin_S_Japan	Leaf	0.12±0.01	0.12	2.16±0.36	2.05
		Stem	0.07±0	0.07	2.26±0.14	2.28
Bors_TY	Sin_Taiwan	Leaf	0.11±0.01	0.11	2.07±0.35	2.03
		Stem	0.1±0.01	0.1	2.02±0.25	2.02

Table F. 6 ABS Lignin in cell wall.

		Plant	ABS Lignin	
Genotype	Genetic group	portion	mean±sd	median
Gawain	Flor_Taiwan	Leaf	23.68±4.51	21.59
		Stem	21.07±3.5	21.47

Geraint	Flor_Taiwan	Leaf	21.98±2.58	20.9
		Stem	21.59±4.62	21.95
Lancelot	Flor_Taiwan	Leaf	24.22±2.85	25.01
		Stem	19.06±2.41	19.06
Percival	Flor_Taiwan	Leaf	21.39±1.59	21.46
		Stem	28.37±2.11	27.79
Lamorak	Hyb.gig	Leaf	18.99±3.94	17.87
		Stem	24.45±1.55	24.68
Galahad	Hyb.mx	Leaf	24.23±3.34	24.72
		Stem	25.59±2.33	25.86
Tristan	Hyb.mx	Leaf	25.06±2.69	24.16
		Stem	23.96±0.76	24.24
Kay	Lut	Leaf	21.21±2.02	21.44
		Stem	23.2±0.92	23.58
Gaheris	Sacc_rob	Leaf	20.22±0.92	19.84
		Stem	24.08±1.19	24.38
Gareth	Sin_N_Japan	Leaf	23.72±1.13	23.76
		Stem	23.93±3.65	21.82
Bedivere	Sin_S_Japan	Leaf	23.3±4.3	21.67
		Stem	23.01±2.05	23.01
Bors_TY	Sin_Taiwan	Leaf	22.4±0.82	22.25
		Stem	22.28±1.44	21.65

Table F. 7 Lignin composition.

Genotype	Genetic group	Plant portion	G		H		S	
			mean±sd	median	mean±sd	median	mean±sd	median
Gawain	Flor_Taiwan	Leaf	54.1±2.08	55.16	4.38±0.25	4.48	41.52±2.33	40.37
		Stem	47.13±1.11	46.82	3.27±0.56	3.54	49.61±0.57	49.64
Geraint	Flor_Taiwan	Leaf	53.73±11.04	47.53	4.32±0.39	4.37	41.95±10.69	47.78
		Stem	45.47±15.48	36.69	4.1±1.3	3.45	50.44±14.78	58.03
Lancelot	Flor_Taiwan	Leaf	59.93±9.93	54.64	4.43±0.64	4.3	35.64±10.55	41.06
		Stem	49.11±6.37	48.47	2.79±0.54	2.74	48.09±6.57	48.17
Percival	Flor_Taiwan	Leaf	51.82±1.14	52.45	3.41±0.56	3.09	44.77±1.49	44.42
		Stem	41.57±2.32	42.03	2.52±0.72	2.29	55.91±2.99	55.68
Lamorak	Hyb.gig	Leaf	68.07±2.84	68.33	2.77±0.13	2.83	29.16±2.82	29.05
		Stem	51.96±0.94	51.61	1.66±0.1	1.72	46.38±0.97	46.86
Kay	Lut	Leaf	73.24±5.94	76.54	3.2±0.87	3	23.55±6.19	21

		Stem	52.75±0.46	52.75	2.52±0.05	2.54	44.73±0.5	44.71
Gaheris	Sacc_rob	Leaf	69.66±1.85	69.54	2.86±0.33	2.86	27.49±1.52	27.6
		Stem	56.49±2.72	57.98	2.61±0.28	2.56	40.9±2.46	39.65
Gareth	Sin_N_Japan	Leaf	63.36±0.09	63.36	4.58±1.41	4.58	32.06±1.5	32.06
		Stem	51.08±1.62	51.46	2.66±0.27	2.68	46.26±1.89	45.86
Bedivere	Sin_S_Japan	Leaf	68.55±3.1	67.99	3.18±0.38	3.08	28.27±3.41	29.14
		Stem	54.97±1.36	55.57	2.45±0.39	2.55	42.58±1.15	42.4

Table F. 8 Lignin architercture traits

Genotype	Genetic group	Plant portion	S:G		H:G	
			mean±sd	median	mean±sd	median
Gawain	Flor_Taiwan	Leaf	0.77±0.07	0.73	0.08±0	0.08
		Stem	1.06±0.03	1.07	0.07±0.01	0.08
Geraint	Flor_Taiwan	Leaf	0.83±0.33	1.01	0.08±0.02	0.09
		Stem	1.26±0.63	1.6	0.1±0.05	0.09
Lancelot	Flor_Taiwan	Leaf	0.62±0.26	0.75	0.07±0	0.07
		Stem	1±0.26	0.99	0.06±0.01	0.05
Percival	Flor_Taiwan	Leaf	0.87±0.05	0.85	0.07±0.01	0.06
		Stem	1.36±0.15	1.34	0.06±0.01	0.06
Lamorak	Hyb.gig	Leaf	0.43±0.06	0.43	0.04±0	0.04
		Stem	0.89±0.03	0.91	0.03±0	0.03
Kay	Lut	Leaf	0.34±0.13	0.28	0.04±0.01	0.04
		Stem	0.85±0.02	0.85	0.05±0	0.05
Gaheris	Sacc_rob	Leaf	0.4±0.04	0.4	0.04±0.01	0.04
		Stem	0.74±0.09	0.68	0.05±0.01	0.04
Gareth	Sin_N_Japan	Leaf	0.51±0.02	0.51	0.07±0.02	0.07
		Stem	0.91±0.07	0.89	0.05±0	0.05
Bedivere	Sin_S_Japan	Leaf	0.41±0.07	0.43	0.05±0	0.05
		Stem	0.78±0.04	0.76	0.04±0.01	0.05

F2 ANOVA tables

Table F. 9 ANOVA test for the effect of genotype andplant portion on monosaccharides enzymatically released.

Variate	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power
GlcE	Plant portion	1	1997.123	1997.123	154.824	0	0.622	0.767	0.615	0.684	0.618	1.815	1
	Genotype	11	276.191	25.108	1.946	0.057	0.086	0.313	0.042	0.128	0.042	0.675	0.913
	Interaction	11	332.626	30.239	2.344	0.021	0.104	0.354	0.059	0.172	0.059	0.741	0.961
	Residuals	47	606.268	12.899									
XylE	Plant portion	1	9.334	9.334	13.345	0.001	0.094	0.221	0.086	0.148	0.087	0.533	0.955
	Genotype	11	29.897	2.718	3.886	0.001	0.3	0.476	0.221	0.309	0.223	0.954	0.999
	Interaction	11	27.586	2.508	3.585	0.001	0.277	0.456	0.198	0.286	0.2	0.916	0.998
	Residuals	47	32.874	0.699									

Table F. 10 Total matrix monosaccharide content in the cell wall

The differences between genotypes and plant portions were tested using ANOVA test.

Variate	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power
Ara	Plant portion	1	20.635	20.635	69.522	0	0.461	0.597	0.451	0.491	0.454	1.216	1
	Genotype	11	7.385	0.671	2.262	0.026	0.165	0.346	0.091	0.164	0.092	0.728	0.954
	Interaction	11	2.833	0.258	0.868	0.577	0.063	0.169	-0.01	-0.021	-0.01	0.451	0.516
	Residuals	47	13.95	0.297									
Fuc	Plant portion	1	0.145	0.145	0.323	0.573	0.004	0.007	-0.009	-0.01	-0.009	0.084	0.088
	Genotype	11	8.778	0.798	1.777	0.086	0.251	0.298	0.108	0.109	0.11	0.652	0.881
	Interaction	11	5.408	0.492	1.095	0.386	0.155	0.208	0.013	0.015	0.013	0.512	0.64
	Residuals	46	20.654	0.449									

Variate	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power
Gal	Plant portion	1	2.815	2.815	81.294	0	0.472	0.634	0.463	0.531	0.466	1.315	1
	Genotype	11	1.1	0.1	2.886	0.006	0.184	0.403	0.12	0.226	0.12	0.822	0.988
	Interaction	11	0.426	0.039	1.118	0.369	0.071	0.207	0.007	0.018	0.008	0.512	0.651
	Residuals	47	1.628	0.035									
Man	Plant portion	1	0.015	0.015	0.012	0.914	0	0	-0.014	-0.014	-0.014	0.016	0.051
	Genotype	11	17.955	1.632	1.255	0.28	0.201	0.227	0.04	0.038	0.041	0.542	0.714
	Interaction	11	10.091	0.917	0.705	0.727	0.113	0.142	-0.047	-0.048	-0.047	0.406	0.419
	Residuals	47	61.128	1.301									
Rha	Plant portion	1	0.022	0.022	0.758	0.389	0.012	0.017	-0.004	-0.004	-0.004	0.133	0.14

Variate	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power
	Genotype	11	0.3	0.027	0.961	0.495	0.168	0.197	-0.007	-0.007	-0.007	0.496	0.57
	Interaction	10	0.248	0.025	0.873	0.565	0.138	0.169	-0.02	-0.02	-0.02	0.45	0.489
	Residuals	43	1.221	0.028									
Xyl	Plant portion	1	7.881	7.881	2.991	0.09	0.029	0.06	0.019	0.027	0.02	0.252	0.409
	Genotype	11	70.588	6.417	2.435	0.017	0.264	0.363	0.154	0.182	0.156	0.755	0.968
	Interaction	11	64.974	5.907	2.242	0.028	0.243	0.344	0.133	0.161	0.135	0.724	0.952
	Residuals	47	123.839	2.635									

Table F. 11 Total cellulose content in the cell wall

Variate	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power

Cel	Plant portion	1	269.098	269.098	67.758	0	0.404	0.59	0.396	0.485	0.398	1.201	1
	Genotype	11	158.77	14.434	3.634	0.001	0.238	0.46	0.172	0.29	0.173	0.922	0.998
	Interaction	11	51.265	4.66	1.173	0.331	0.077	0.215	0.011	0.026	0.011	0.524	0.677
	Residuals	47	186.658	3.971									
Hem	Plant portion	1	0.104	0.104	0.013	0.911	0	0	-0.015	-0.014	-0.015	0.016	0.051
	Genotype	11	66.511	6.046	0.738	0.697	0.126	0.145	-0.044	-0.042	-0.045	0.411	0.438
	Interaction	11	68.036	6.185	0.755	0.682	0.129	0.147	-0.041	-0.039	-0.042	0.416	0.449
	Residuals	48	393.441	8.197									

Table F. 12 Cell wall architecture

Variate	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power

Ara/Xyl	Plant portion	1	0.036	0.036	108.72	0	0.526	0.698	0.519	0.603	0.521	1.521	1
	Genotype	11	0.012	0.001	3.162	0.003	0.168	0.425	0.115	0.251	0.115	0.86	0.994
	Interaction	11	0.005	0	1.465	0.177	0.078	0.255	0.025	0.067	0.025	0.586	0.795
	Residuals	47	0.016	0									
Cel/Hem	Plant portion	1	0.389	0.389	5.5	0.023	0.063	0.105	0.051	0.06	0.051	0.342	0.65
	Genotype	11	1.907	0.173	2.453	0.016	0.308	0.365	0.181	0.184	0.183	0.758	0.969
	Interaction	11	0.565	0.051	0.727	0.707	0.091	0.145	-0.034	-0.044	-0.034	0.413	0.432
	Residuals	47	3.321	0.071									

Table F. 13 ABS Lignin. Total content in the cell wall.

Variate	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power

	Plant												
ABSL	potion	1	16.793	16.793	2.347	0.132	0.027	0.048	0.015	0.019	0.015	0.223	0.335
	Genotype	11	95.45	8.677	1.213	0.305	0.152	0.221	0.026	0.032	0.027	0.533	0.695
	Interaction	11	177.424	16.129	2.254	0.027	0.283	0.345	0.156	0.163	0.158	0.726	0.953
	Residuals	47	336.306	7.155									

Table F. 14 Lignin composition.

Quantification of the three main monolignols in the cell wall of leaf and stem of 12 *Miscanthus* genotypes

Variate	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power
			2038.06	2038.06									
G	Plant portion	1	9	9	61.252	0	0.388	0.636	0.379	0.532	0.382	1.323	1
	Genotype	8	1846.569	230.821	6.937	0	0.352	0.613	0.299	0.473	0.301	1.259	1

									-	-		
	Interaction	8	202.275	25.284	0.76	0.64	0.039	0.148	0.012	-0.038	0.012	0.417	0.37
	Residuals	35	1164.574	33.274									
H	Plant portion	1	11.102	11.102	31.95	0	0.242	0.477	0.232	0.369	0.234	0.955	1
	Genotype	8	19.064	2.383	6.858	0	0.415	0.611	0.352	0.469	0.354	1.252	1
	Interaction	8	3.621	0.453	1.303	0.274	0.079	0.229	0.018	0.044	0.018	0.546	0.62
	Residuals	35	12.162	0.347									
S	Plant portion	1	2350.017	2350.017	71.218	0	0.44	0.67	0.432	0.57	0.434	1.426	1
	Genotype	8	1628.829	203.604	6.17	0	0.305	0.585	0.254	0.438	0.256	1.188	1
	Interaction	8	202.07	25.259	0.765	0.635	0.038	0.149	- 0.012	-0.037	- 0.012	0.418	0.373
	Residuals	35	1154.918	32.998									
S/G	Plant portion	1	2.157	2.157	53.147	0	0.383	0.603	0.373	0.496	0.376	1.232	1
	Genotype	8	1.99	0.249	6.13	0	0.353	0.584	0.294	0.436	0.296	1.184	1

									-		-		
	Interaction	8	0.063	0.008	0.194	0.99	0.011	0.043	0.046	-0.138	0.046	0.211	0.112
	Residuals	35	1.42	0.041									
						0.38	0.00		-		-		
H/G	Plant portion	1	0	0	0.776	4	8	0.022	0.002	-0.004	0.002	0.149	0.143
	Genotype	8	0.015	0.002	7.406	0	0.585	0.629	0.501	0.492	0.506	1.301	1
									-		-		
	Interaction	8	0.002	0	0.772	0.63	0.061	0.15	0.018	-0.036	0.018	0.42	0.376
	Residuals	35	0.009	0									

F3 Pictures

Genotype 3 - Lancelot

Rep 1 Rep 2 Rep 3

Genotype 8 – Gawain

Rep 1

Rep 2

Rep 3

Genotype 9 – Geraint

Rep 1

Rep 2

Rep 3

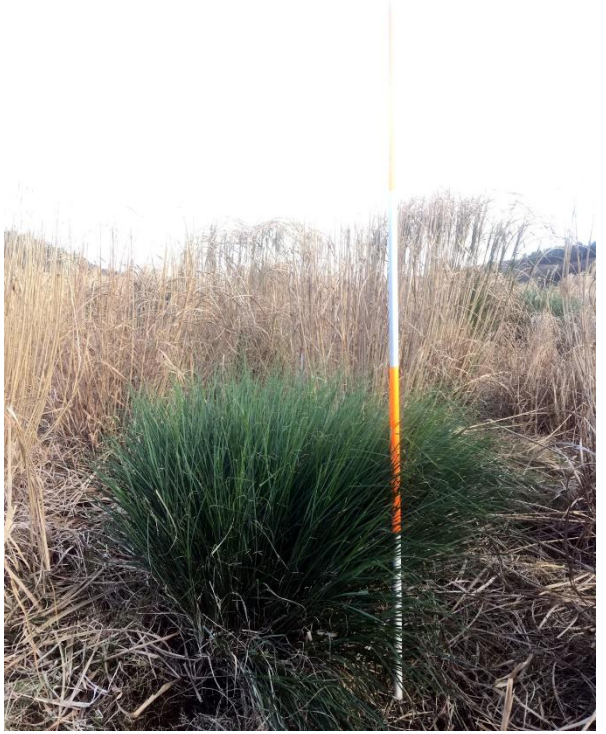
Genotype 10 – Percival

Rep 1

Rep 2

NA

Rep 3



Genotype 18 - Bors

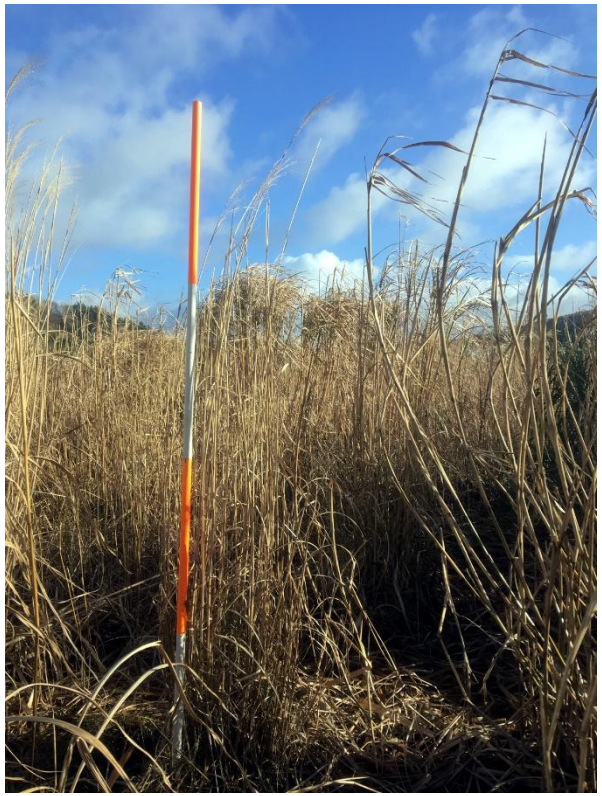
Rep 1

Rep 2

Rep 3

Genotype 20 - Lamorak

Rep 1



Rep 2

Rep 3

Genotype 30 – Kay

Rep 1 Rep 2 Rep 3

Genotype 38 - Gareth

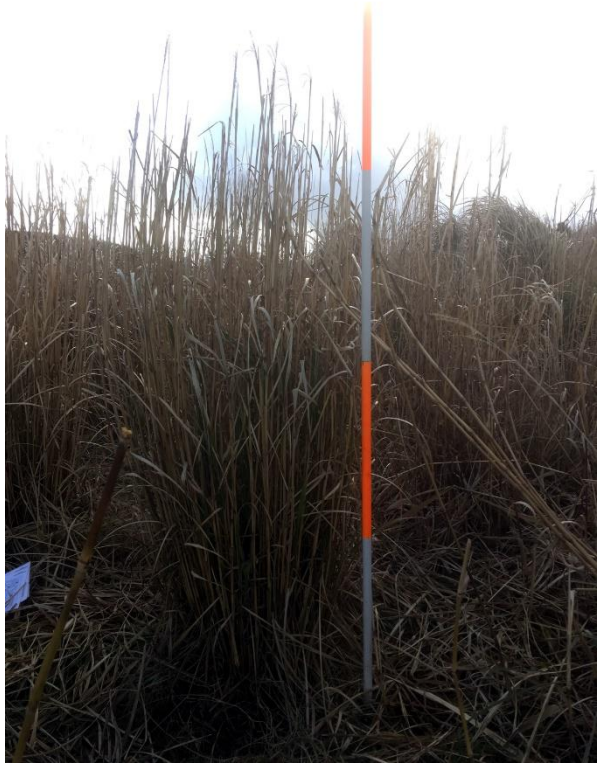
Rep 1

Rep 2

Rep 3

Genotype 39 - Bedivere

Rep 1



Rep 2

Rep 3

Genotype 41 – Gaheris

Rep 1

Rep 2

Rep 3

Genotype 48 - Galahd

Rep 1

Rep 2

Rep 3

Genotype 49 - Tristan

Rep 1

Rep 2

Rep 3

Appendix G Chapter 6

G1 Tables

Parameter	Description
t_for_Fm	Time to Fm
Area	Area under curve
Fo	Minimum value for chlorophyll fluorescence
Fm	Maximum value for chlorophyll fluorescence
Fv	Variable fluorescence
F1	Value of fluorescence at O
F2	Value of fluorescence at J
F3	Value of fluorescence at I
F4	Value of fluorescence at P
F5	
Fo.Fm	Ratio between the extremes expresses the amplitude of the ChlF transient rise
Fv.Fm	Maximum quantum efficiency of PSII photochemistry.
PI_Inst.	Instantaeous Photosynthetic Index
Fv.Fo	Ratio Fv/F0
Vj	Relative variable fluorescence at the J-step
Vi	Relative variable fluorescence at the I-step
dVG.dto	dVG.dto
dV.dto	dV.dto
Sm	Normalized area above the transient.
N	Number of turnovers of QA
Sm.t(Fm)	Normalized area above the transient at Fm
ABS.RC	Absorption flux (of antenna Chls) per RC
DIo.RC	Dissipated energy flux per RC (at t = 0)
TRo.RC	Trapping flux (leading to QA reduction) per RC
ETo.RC	Electron transport flux (further than) per RC
REo.RC	Electron flux reducing end electron acceptors at the PSI acceptor side per RC
phi(Po)	Maximum quantum yield of primary photochemistry (at t = 0)
psi(Eo)	Probability an exciton moves an electron beyond Qa
phi(Eo)	Quantum yield of electron transport (at t = 0)
delta(Ro)	Prob.that an electron from the intersystem reduces electron acceptors at the PSI end
phi(Ro)	Quantum yield for reduction of end electron acceptors at the PSI acceptor side
ABS.CSo	Absorption flux for cross section at 0
DIo.CSo	Dissipated energy flux per cross section at 0
TRo.CSo	Trapped energyflux (leading to QA reduction) per cross section at 0
ETo.CSo	Electron transportflux per cross sectionat 0
REo.CSo	Electronflux reducing end electron acceptors at the PSI, per cross section at 0
ABS.CSm	Absorption flux for cross section at Fm
DIo.CSm	Dissipated energy flux per cross section at at Fm
TRo.CSm	Trapped energyflux (leading to QA reduction) per cross section at at Fm
ETo.CSm	Electron transportflux per cross sectionat at Fm
REo.CSm	Electronflux reducing end electron acceptors at the PSI cross section at at Fm
gamma	gamma
phi(Po)_1	aximum quantum yield for primaryphotochemistry
psi(Eo)_1	Probability (at t = 0) that a trapped exciton moves an electron QA
PI_abs	Performance index
dRo	Probability that an electron is transferred to reduce endelectron acceptors at the PSI
PI_total	Performance index
DF_abs	Delayed (chlorophyll) fluorescence Absolute
DF_Total	Delayed (chlorophyll) fluorescence total
kPABSkF	kPABSkF
kNABSkF	kNABSkF
Origin_to_F1	Origin to F1
F1_to_F3	F1 to F3
F1_to_F4	F1 to F4
F1_to_F5	F1 to F5
F3_to_F4	F3 to F4
F4_to_F5	F4 to F5
F5_to_Fm	F5 to Fm

Figure S 1 List of photosynthetic parameters measured by the HandyPea apparatus.

G2Summary statistics

Table G. 1 Dry weight of plant, stem and leaf and relative leaf content

Genetic group	Genotype	Treatment	Plant DW (g)		Stem DW (g)		Leaf DW (g)		Leaf (%)	
			mean±sd	median	mean±sd	median	mean±sd	median	mean±sd	median
Flor 2	Gawain	Control	4.63±6.88	1.3	0.42±0.6	0.42	4.35±6.4	1.3	0.98±0.04	1
		Minus N	1.03±0.73	1.25	0.02±0.04	0	1±0.72	1.25	0.97±0.04	0.98
		Minus P	6.4±NA	6.4	0±NA	0	6.4±NA	6.4	1±NA	1
	Geraint	Control	5.25±4.43	5.85	0±NA	0	5.25±4.43	5.85	1±0	1
		Minus N	0.42±0.3	0.45	0±0	0	0.42±0.3	0.45	1±0	1
		Minus P	3.8±3.6	2.35	0±0	0	3.8±3.6	2.35	1±0	1
	Lancelot	Control	3.97±6.09	0.6	0.6±1.04	0	3.37±5.05	0.6	0.95±0.09	1
		Minus N	0.62±0.9	0.17	0.06±0.1	0	0.6±0.87	0.17	0.99±0.02	1

Genetic group	Genotype	Treatment	Plant DW (g)		Stem DW (g)		Leaf DW (g)		Leaf (%)	
			mean±sd	median	mean±sd	median	mean±sd	median	mean±sd	median
		Minus P	3.75±4.31	3.75	0.95±1.34	0.95	2.8±2.97	2.8	0.86±0.2	0.86
	Percival	Control	0.65±0.16	0.68	0±0	0	0.65±0.16	0.68	1±0	1
		Minus N	0.1±0.07	0.12	0±0	0	0.14±0.03	0.14	1±0	1
		Minus P	0.91±0.38	1	0±0	0	0.91±0.38	1	1±0	1
Hyb.gig	Lamorak	Control	18.31±19.71	7.85	5.43±6.52	1.35	12.88±13.23	6.5	0.8±0.13	0.8
		Minus N	1.55±1.67	1.15	0.58±0.84	0.2	1.48±0.77	1.2	0.82±0.2	0.86
		Minus P	10.94±10.88	6.85	2.95±3.23	1.73	7.99±7.66	5.12	0.74±0.04	0.72
Hyb.mx	Galahad	Control	20.1±9.37	18.7	3.65±1.94	3.7	16.45±7.45	15	0.83±0.03	0.81
		Minus N	0.9±0.44	0.82	0.15±0.15	0.12	0.75±0.29	0.7	0.87±0.11	0.86
		Minus P	3.46±1.87	4.1	0.8±0.48	0.9	2.66±1.4	3.2	0.81±0.11	0.77
	Tristan	Control	21.82±12.39	23	4.99±3.39	5.15	16.83±9.06	17.85	0.79±0.07	0.76
		Minus N	0.98±0.59	0.73	0.13±0.04	0.13	0.85±0.61	0.6	0.83±0.09	0.82
		Minus P	7.2±6.63	5.8	1.9±1.05	2.05	5.68±5.42	4.15	0.81±0.11	0.8

Genetic group	Genotype	Treatment	Plant DW (g)		Stem DW (g)		Leaf DW (g)		Leaf (%)	
			mean±sd	median	mean±sd	median	mean±sd	median	mean±sd	median
Lut	Kay	Control	13.03±17.08	13.03	3.98±5.62	3.98	9.05±11.46	9.05	0.84±0.23	0.84
		Minus N	1.59±1.38	1.35	0.67±0.63	0.75	1.09±0.78	0.98	0.82±0.21	0.82
		Minus P	3.38±4.77	3.38	2.95±NA	2.95	3.8±NA	3.8	0.56±NA	0.56
Sacc_rob	Gaheris	Control	19.26±10.98	20.8	4.02±2.49	4.15	15.24±8.54	17.15	0.82±0.07	0.79
		Minus N	0.71±0.31	0.85	0.31±0.16	0.38	0.41±0.17	0.45	0.59±0.08	0.58
		Minus P	5.14±6.69	1.25	1.34±1.53	1.1	4.07±5.25	1.1	0.89±0.11	0.88
Sin_N_Japan	Gareth	Control	3.08±4.21	1.48	0.16±0.29	0.02	3.88±4.18	1.65	0.98±0.03	1
		Minus N	0.38±0.43	0.27	0.02±0.03	0	0.36±0.4	0.27	0.97±0.04	0.97
		Minus P	3.42±4.64	3.42	0.18±0.25	0.18	3.24±4.4	3.24	0.98±0.04	0.98
Sin_S_Japan	Bedivere	Control	10.14±15.65	3.75	3.49±6.36	0.48	6.65±9.41	3.27	0.87±0.18	0.94
		Minus N	3.28±4.56	0.75	1.78±3.09	0	1.5±1.48	0.75	0.79±0.36	1
		Minus P	5.66±10.07	0.84	1.82±3.55	0.07	3.84±6.52	0.77	0.87±0.15	0.91
Sin_Taiwan	Bors t.Y.	Control	6.51±6.92	4.6	1.62±2.2	0.15	4.89±4.87	4.45	0.86±0.17	0.96

Genetic group	Genotype	Treatment	Plant DW (g)		Stem DW (g)		Leaf DW (g)		Leaf (%)	
			mean±sd	median	mean±sd	median	mean±sd	median	mean±sd	median
		Minus N	1.86±1.07	2	0.33±0.32	0.15	1.53±1.07	1.3	0.81±0.15	0.82
		Minus P	5.58±3.46	4.3	1±0.7	0.7	4.58±2.85	3.6	0.83±0.06	0.84

Table G. 2 Dry weight of the plant, single tiller and number of tillers

Genetic group	Genotype	Treatment	Plant DW (g)		Tillers (n)		Tiller DW (g/tiller)	
			mean±sd	median	mean±sd	median	mean±sd	median
Flor	Gawain	Control	4.63±6.88	1.3	4±3	4	1.03±0.92	1.3
		Minus N	1.03±0.73	1.25	1.25±0.5	1	1.08±0.57	1.02
		Minus P	6.4±NA	6.4	2±NA	2	3.2±NA	3.2
	Geraint	Control	5.25±4.43	5.85	13.67±6.66	17	0.32±0.22	0.34
		Minus N	0.42±0.3	0.45	2±0.82	2	0.19±0.1	0.21
		Minus P	3.8±3.6	2.35	10.75±4.57	10.5	0.35±0.25	0.28

Genetic group	Genotype	Treatment	Plant DW (g)		Tillers (n)		Tiller DW (g/tiller)	
			mean±sd	median	mean±sd	median	mean±sd	median
	Lancelot	Control	3.97±6.09	0.6	2.67±1.53	3	1.08±1.44	0.3
		Minus N	0.62±0.9	0.17	1.33±0.58	1	0.59±0.92	0.09
		Minus P	3.75±4.31	3.75	1.5±0.71	1.5	2.05±1.91	2.05
	Percival	Control	0.65±0.16	0.68	10.25±2.06	10.5	0.06±0.02	0.06
		Minus N	0.1±0.07	0.12	3±0.82	3	0.04±0.03	0.04
		Minus P	0.91±0.38	1	8±3.39	9	0.12±0.01	0.11
Hyb.gig	Lamorak	Control	18.31±19.71	7.85	6±2	6	2.46±2.32	1.31
		Minus N	1.55±1.67	1.15	1.75±1.5	1	1.55±1.67	1.15
		Minus P	10.94±10.88	6.85	3±2.65	2	4.56±3.06	4.5
Hyb.mx	Galahad	Control	20.1±9.37	18.7	9±2.74	9	2.21±0.8	2.39
		Minus N	0.9a□0.44	0.82	1±0	1	0.9±0.44	0.82
		Minus P	3.46±1.87	4.1	2±1.73	1	2.59±1.89	3.1
	Tristan	Control	21.82±12.39	23	9.4±2.3	9	2.27±1.1	2.32

Genetic group	Genotype	Treatment	Plant DW (g)		Tillers (n)		Tiller DW (g/tiller)	
			mean±sd	median	mean±sd	median	mean±sd	median
		Minus N	0.98±0.59	0.73	1.5±0.58	1.5	0.67±0.27	0.72
		Minus P	7.2±6.63	5.8	4.6±3.05	5	1.41±0.86	1.78
Lut	Kay	Control	13.03±17.08	13.03	7.5±2.12	7.5	2.15±2.88	2.15
		Minus N	1.59±1.38	1.35	1±0	1	1.59±1.38	1.35
		Minus P	3.38±4.77	3.38	2.5±2.12	2.5	3.38±4.77	3.38
Sacc_rob	Gaheris	Control	19.26±10.98	20.8	11.4±6.19	13	1.56±0.6	1.51
		Minus N	0.71±0.31	0.85	1±0	1	0.71±0.31	0.85
		Minus P	5.14±6.69	1.25	4.4±3.29	2	0.76±0.77	0.63
Sin_N_Japan	Gareth	Control	3.08±4.21	1.48	2.33±1.15	3	1.65±1.31	1.3
		Minus N	0.38±0.43	0.27	1±0	1	0.44±0.59	0.44
		Minus P	3.42±4.64	3.42	4.5±0.71	4.5	0.68±0.93	0.68
Sin_S_Japan	Bedivere	Control	10.14±15.65	3.75	4±2.94	3.5	1.53±1.92	0.99
		Minus N	3.28±4.56	0.75	1±0	1	3.28±4.56	0.75

Genetic group	Genotype	Treatment	Plant DW (g)		Tillers (n)		Tiller DW (g/tiller)	
			mean±sd	median	mean±sd	median	mean±sd	median
Sin_Taiwan	Bors t.Y.	Minus P	5.66±10.07	0.84	3.5±3	3	1.25±1.94	0.32
		Control	6.51±6.92	4.6	2.8±1.64	2	2.16±1.72	1.77
		Minus N	1.86±1.07	2	1.67±1.15	1	1.22±0.68	0.95
		Minus P	5.58±3.46	4.3	3.2±1.48	3	2.08±1.39	1.9

Table G. 3 Result of the enzymatic monosaccharide release from cell wall.

Cell wall material was digested for 72 hours with a cell wall degrading enzymatic mix (CellTech, Novozyme). At the end of the digestion the amount of arabinose (AraE), glucose (GlcE) and xylose (XylE) released in the supernatant were measured. The values were obtained averaging the result of 3 technical replicates. Values represent the mean ± standard deviation of the mean (sd) and median between the values obtained in the three biological replicates.

Genotype	Genetic group	Treatment	AraE		GlcE		XylE	
			mean±sd	median	mean±sd	median	mean±sd	median

Bors the Younger	Sin_Taiwan	C	6.619±1.596	6.955	14.148±2.434	13.992	3.413±1.091	3.764
		N-	5.452±1.119	5.695	15.228±4.483	14.612	2.958±0.958	2.742
		P-	4.261±1.056	4.368	14.123±1.738	13.118	2.124±0.763	2.148
Gaheris	SaccRob	C	6.303±2.914	6.345	14.214±13.03	8.599	3.025±1.172	2.269
		N-	13.181±0.853	13.666	43.924±12.285	39.128	5.781±2.07	4.656
		P-	17.725±8.844	14.399	57.375±22.773	48.068	11.285±6.951	9.902
Galahad	Hyb.mx	C	14.273±0.975	14.033	45.147±5.891	47.418	10.28±2.774	10.735
		N-	8.075±3.071	7.724	23.172±9.924	23.335	4.118f1.557	4.11
		P-	10.387±5.125	7.577	28.714±14.863	27.884	5.324±2.226	4.596
Lamorak	Hyb.gig	C	6.021±1.449	5.921	18.224±3.563	18.584	2.973±0.758	2.817
		N-	6.281±1.324	6.129	19.015±5.692	19.109	3.711±1.406	3.892
		P-	4.798±1.365	4.648	11.8±NA	11.8	2.228±0.803	2.255
Tristan	Hyb.mx	C	18.384±4.707	19.635	42.581±16.473	38.999	7.767±2.808	6.817
		N-	17.432±6.102	18.206	50.605±3.882	49.299	6.949±2.501	7.231
		P-	9.505±1.714	9.991	16.583±2.792	16.147	3.5±0.68	3.832

Table G. 4 Matrix monosaccharides content – Part 1

Genotype	Genetic group	Treatment	Fuc		Ara		Glc	
			mean±sd	median	mean±sd	median	mean±sd	median
Bors the Younger	Sin_Taiwan	C	0.38±0.07	0.404	7.241±0.88	6.987	6.087±1.038	5.927
		N-	0.454±0.176	0.423	8.108±1.091	8.221	13.878±7.199	10.096
		P-	0.446±0.16	0.434	7.851±0.805	7.61	8.246±2.397	9.258
Gaheris	SaccRob	C	0.65±0.185	0.553	5.21±1.932	5.916	8.986±5.683	8.327
		N-	0.494±0.081	0.475	5.434±0.567	5.229	8.878±1.967	9.976
		P-	0.58±0.312	0.508	4.847±0.836	4.879	7.184±0.838	6.846
Galahad	Hyb.mx	C	0.789±0.226	0.794	5.405±0.971	5.021	8.232±1.076	8.045
		N-	0.718±0.176	0.721	5.751±1.194	5.735	9.53±2.385	9.07
		P-	0.772±0.176	0.719	5.064±0.945	4.79	10.237±0.342	10.321
Lamorak	Hyb.gig	C	0.595±0.077	0.602	7.08±0.568	7.287	5.324±1.014	4.972
		N-	0.469±0.186	0.435	7.151±0.319	7.102	11.714±7.192	8.805
		P-	0.652±0.187	0.668	5.788±1.136	6.029	6.449±1.442	6.144

Tristan	Hyb.mx	C	0.58±0.179	0.639	3.796±0.903	3.758	4.631±1.077	4.689
		N-	0.805±0.137	0.767	4.217±1.282	4.041	9.677±2.475	10.823
		P-	0.844±0.252	0.79	3.61±0.645	3.495	4.851±1.216	4.894

Table G. 5 Matrix monosaccharides content – Part 2

Genotype	Genetic group	Treatment	Xyl		Man		Cel	
			mean±sd	median	mean±sd	median	mean±sd	median
Bors the Younger	Sin_Taiwan	C	40.946±6.661	38.287	0.927±0.18	0.934	0.645±0.19	0.723
		N-	41.286±7.042	40.273	0.751±0.322	0.657	0.853±0.255	0.766
		P-	44.251±4.512	44.33	1.046±0.117	1.052	0.682±0.241	0.811
Gaheris	SaccRob	C	24.588±6.188	21.687	1.498±1.891	0.761	0.607±0.261	0.481
		N-	25.67±3.855	26.071	0.249±0.144	0.211	0.575±0.148	0.54
		P-	21.436±9.316	18.63	0.518±0.308	0.454	0.579±0.38	0.456

Genotype	Genetic group	Treatment	Xyl		Man		Cel	
			mean±sd	median	mean±sd	median	mean±sd	median
Galahad	Hyb.mx	C	16.563±2.318	16.127	0.754±0.146	0.765	0.502±0.092	0.519
		N-	23.628±5.775	24.486	0.317±0.142	0.286	0.571±0.357	0.443
		P-	23.763±2.577	22.536	0.694±0.496	0.711	0.68±0.273	0.749
Lamorak	Hyb.gig	C	35.787±4.066	35.389	1.041±0.322	1.051	0.797±0.107	0.845
		N-	36.85±7.041	38.102	0.848±0.387	0.765	0.91±0.171	0.945
		P-	30.022±4.112	30.597	0.753±0.104	0.751	0.738±0.154	0.781
Tristan	Hyb.mx	C	23.936±5.003	25.662	0.795±0.153	0.769	0.501±0.267	0.471
		N-	24.439±1.458	24.364	0.532±0.093	0.528	0.8±0.195	0.838
		P-	23.283±5.129	21.969	0.744±0.275	0.748	0.814±0.376	0.8

Table G. 6

Genotype	Treatment	CrCel	Hem	MnsT

	Genetic group		mean±sd	median	mean±sd	median	mean±sd	median
Bors the Younger	Sin_Taiwan	C	46.119±6.258	46.044	55.581±8.377	51.979	88.759±8.733	89.766
		N-	38.382±8.628	40.179	64.353±5.403	67.438	97.367±17.164	99.72
		P-	36.411±7.145	38.355	61.557±7.312	60.966	87.334±9.243	90.896
Gaheris	SaccRob	C	38.356±6.542	35.353	40.931±6.233	42.891	68.051±5.635	68.015
		N-	30.865±11.985	32.144	40.697±2.824	41.905	64.299±1.897	64.155
		P-	21.249±2.167	21.374	28.398±8.047	31.385	41.156±11.055	41.945
Galahad	Hyb.mx	C	22.247±3.264	21.496	31.744±3.125	31.343	43.761±2.843	44.188
		N-	35.462±15.336	29.192	39.873±7.36	41.648	62.036±11.497	62.917
		P-	27.141±7.733	27.252	40.531±3.519	40.695	58.549±7.871	58.958
Lamorak	Hyb.gig	C	33.325±2.304	33.261	49.826±5.412	50.156	71.853±6.454	72.879
		N-	38.228±7.86	35.457	56.892±4.039	57.844	84.564±9.944	81.369
		P-	42.244±NA	42.244	43.663±5.431	45.076	71.443±6.851	74.455
Tristan	Hyb.mx	C	22.941±4.353	22.637	33.738±6.649	36.362	48.125±7.462	49.783

N-	24.091±3.228	22.311	39.595±4.479	39.625	59.288±6.176	58.001
P-	28.383±6.01	29.329	33.332±6.749	30.805	60.312±20.362	50.909

Table G. 7 Cell wall architecture traits

Genotype	Genetic group	Treatment	Ara/Xyl		Cel/Hem	
			mean±sd	median	mean±sd	median
Bors the Younger	Sin_Taiwan	C	0.178±0.009	0.177	0.986±0.167	1.009
		N-	0.272±0.176	0.172	0.777±0.082	0.778
		P-	0.178±0.005	0.179	0.782±0.191	0.761
Gaheris	SaccRob	C	0.264±0.153	0.275	1.203±0.228	1.145
		N-	0.224±0.036	0.244	0.977±0.099	0.95
		P-	0.251±0.058	0.268	1.004±0.097	0.982
Galahad	Hyb.mx	C	0.354±0.071	0.336	1.004±0.118	0.956
		N-	0.27±0.021	0.275	1.064±0.109	1.095
		P-	0.223±0.041	0.232	1.046±0.23	0.978
Lamorak	Hyb.gig	C	0.2±0.022	0.205	0.772±0.056	0.777
		N-	0.202±0.053	0.18	0.882±0.046	0.907
		P-	0.195±0.028	0.184	1.214±NA	1.214
Tristan	Hyb.mx	C	0.171±0.039	0.186	0.981±0.288	0.905
		N-	0.179±0.034	0.176	0.927±0.059	0.923
		P-	0.156±0.006	0.158	1.06±0.299	1.122

G3ANOVA tables

Table G. 8 ANOVA table of the effect of genotype and genetic group on stem DW and leaf %

Variate	Source	df	SS	MS	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power
DW (wt/tiller)	Genotype	11	75.345	6.85	3.07	0.001	0.218	0.218	0.146	0.146	0.147	0.528	0.992
	Residuals	121	269.943	2.231									
	Genetic group	7	62.589	8.941	3.954	0.001	0.181	0.181	0.135	0.135	0.135	0.471	0.986
	Residuals	125	282.699	2.262									
	Treatment	2	11.408	5.704	2.221	0.113	0.033	0.033	0.018	0.018	0.018	0.185	0.455
	Residuals	130	333.88	2.568									

Leaf (wt%/tiller)	Genotype	11	0.919	0.084	6.093	0	0.362	0.362	0.301	0.301	0.303	0.754	1
	Residuals	118	1.618	0.014									
	Genetic group	7	0.894	0.128	9.488	0	0.352	0.352	0.314	0.314	0.315	0.738	1
	Residuals	122	1.643	0.013									
	Treatment	2	0.005	0.002	0.115	0.892	0.002	0.002	-0.014	-0.014	-0.014	0.043	0.068
	Residuals	127	2.533	0.02									

Table G. 9 Effect of genotype and treatment on monosaccharides enzymatically released.

Two-way ANOVA test with interaction. The effect of genotype and treatment on the amount of arabinose (AraE), glucose (GlcE) and xylose (XylE) released upon 72 hours digestion with an enzymatic cocktail was tested. As a consequence of the significantinteraction the result cannot be interpreted.

Variate 1	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
							0.00		-		-		
AraE	Treatment	12.4	6.2	2	0.45	0.64	7	0.019	0.008	-0.018	0.008	0.138	0.124
	Genotype	615.563	153.891	4	11.17	0	0.331	0.487	0.299	0.396	0.301	0.975	1
	Interactio												
	n	584.093	73.012	8	5.299	0	0.314	0.474	0.253	0.357	0.255	0.95	1
	Residuals	647.534	13.777	47									
						0.98							
GlcE	Treatment	2.695	1.347	2	0.011	9	0	0.001	-0.014	-0.035	-0.014	0.023	0.052
		4390.97	1097.74										
	Genotype	9	5	4	9.175	0	0.254	0.46	0.225	0.361	0.227	0.924	1
	Interactio	7726.86					0.44						
	n	2	965.858	8	8.073	0	8	0.6	0.389	0.494	0.392	1.226	1
	Residuals	5144.593	119.642	43									

					0.32		0.00						
XylE	Treatment	3.844	1.922	2	6	0.723	5	0.014	-0.01	-0.022	-0.01	0.118	0.102
	Genotype	196.7	49.175	4	8.35	0	0.258	0.415	0.225	0.322	0.227	0.843	0.999
	Interactio				6.06								
	n	285.946	35.743	8	9	0	0.375	0.508	0.311	0.395	0.313	1.016	1
	Residuals	276.79	5.889	47									

Table G. 10 One Way ANOVAs for the effect of the treatment on the monosaccharide release of the single genotypes

Separate ANOVA tests for the effect of the treatment on each genotype separately (split dataset) were performed on the amount of the three monosaccharides (Mns), arabinose (AraE), glucose (Glc) and xylose (XylE) released after 72-h digestion with a cell degrading enzymatic mix (CellTech)

Mns	Genotype	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's 'f	power
AraE	Bors	Treatment	12.4	6.2	2	3.822	0.063	0.459	0.459	0.32	0.32	0.339	0.922	0.669
		Residuals	14.599	1.622	9									
	Gaheris	Treatment	297.257	148.628	2	4.953	0.035	0.524	0.524	0.397	0.397	0.418	1.049	0.785
		Residuals	270.086	30.01	9									

Mns	Genotype	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's 'f	power
	Galahad	Treatment	78.737	39.369	2	2.89	0.102	0.366	0.366	0.225	0.225	0.24	0.76	0.547
		Residuals	136.226	13.623	10									
	Lamorak	Treatment	4.775	2.387	2	1.228	0.338	0.214	0.214	0.037	0.037	0.04	0.522	0.26
		Residuals	17.496	1.944	9									
	Tristan	Treatment	200.045	100.022	2	4.783	0.035	0.489	0.489	0.368	0.368	0.387	0.978	0.775
		Residuals	209.127	20.913	10									

Mns	Genotype	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's 'f	power
GlcE	Bors	Treatment	2.695	1.347	2	0.173	0.844	0.037	0.037	-0.16	-0.16	-0.177	0.196	0.076
		Residuals	70.052	7.784	9									
	Gaheris	Treatment	4389.12	2194.56	2	7.786	0.011	0.634	0.634	0.531	0.531	0.552	1.315	0.936
		Residuals	2536.763	281.863	9									
	Galahad	Treatment	1057.031	528.516	2	4.119	0.05	0.452	0.452	0.324	0.324	0.342	0.908	0.708
		Residuals	1283.21	128.321	10									

Mns	Genotype	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's 'f	power
XylE	Lamorak	Treatment	41.325	20.663	2	1.073	0.4	0.263	0.263	0.016	0.016	0.018	0.598	0.223
		Residuals	115.583	19.264	6									
	Tristan	Treatment	2364.463	1182.231	2	9.342	0.006	0.675	0.675	0.582	0.582	0.603	1.441	0.969
		Residuals	1138.984	126.554	9									
	Bors	Treatment	3.844	1.922	2	2.238	0.163	0.332	0.332	0.171	0.171	0.184	0.705	0.439
		Residuals	7.73	0.859	9									

Mns	Genotype	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's 'f	power
	Gaheris	Treatment	153.502	76.751	2	4.344	0.048	0.491	0.491	0.358	0.358	0.378	0.983	0.727
		Residuals	159.005	17.667	9									
	Galahad	Treatment	86.753	43.377	2	8.643	0.007	0.634	0.634	0.54	0.54	0.56	1.315	0.959
		Residuals	50.185	5.019	10									
	Lamorak	Treatment	3.804	1.902	2	2.092	0.179	0.317	0.317	0.154	0.154	0.166	0.682	0.414
		Residuals	8.184	0.909	9									

Mns	Genotype	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's 'f	power
	Tristan	Treatment	12.4	6.2	2	3.822	0.063	0.459	0.459	0.32	0.32	0.339	0.922	0.669
		Residuals	14.599	1.622	9									

Table G. 11 One Way ANOVAs for the effect of the treatment on the monosaccharide release of the single genotypes

Separate ANOVA tests for the effect of the genotype in each treatment group separately (split dataset along treatment factor) were performed on the amount of the three monosaccharides (Mns), arabinose (AraE), glucose (Glc) and xylose (XylE) released after 72-h digestion with a cell degrading enzymatic mix (CellTech)

Monosaccharide	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's 'f	power
AraE	C	Genotype	615.563	153.891	4	19.580	0	0.813	0.813	0.764	0.764	0.772	2.086	1

		Residuals	141.476	7.86	18								
		Genotype	3058.512	764.628	4	11.484	0.001	0.807	0.724	0.724	0.737	2.044	1
GlcE	N-	Residuals	732.393	66.581	11								
		Genotype	238.586	59.646	4	5.953	0.003	0.583	0.474	0.474	0.485	1.183	0.983
XylE	P-	Residuals	170.418	10.025	17								

Table G. 12 Two-way ANOVA with interaction for the effect of genotype and treatment on the matrix monosaccharides

Mns	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's 'f	power
Fuc	Treatment	0.013	0.006	2	0.192	0.826	0.006	0.008	-0.024	-0.027	-0.025	0.09	0.08
	Genotype	0.35	0.088	4	2.633	0.046	0.16	0.183	0.098	0.095	0.099	0.473	0.742
	Trt:Gen	0.262	0.033	8	0.986	0.459	0.12	0.144	-0.002	-0.002	-0.002	0.41	0.479
	Residuals	1.563	0.033	47									
Ara	Treatment	1.452	0.726	2	0.677	0.513	0.015	0.028	-0.007	-0.011	-0.007	0.17	0.165
	Genotype	38.698	9.675	4	9.013	0	0.408	0.434	0.359	0.341	0.363	0.876	0.999

Mns	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's 'f	power
	Trt:Gen	4.139	0.517	8	0.482	0.863	0.044	0.076	-0.046	-0.072	-0.047	0.286	0.234
	Residuals	50.449	1.073	47									
Glc	Treatment	108.109	54.054	2	5.892	0.005	0.151	0.2	0.123	0.136	0.125	0.501	0.876
	Genotype	67.7	16.925	4	1.845	0.136	0.094	0.136	0.043	0.052	0.043	0.396	0.566
	Trt:Gen	111.21	13.901	8	1.515	0.178	0.155	0.205	0.052	0.062	0.053	0.508	0.699
	Residuals	431.176	9.174	47									

Mns	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's 'f	power
Xyl	Treatment	29.31	14.655	2	0.529	0.592	0.009	0.022	-0.008	-0.015	0.008	0.15	0.138
	Genotype	1635.08	408.772	4	14.766	0	0.504	0.557	0.466	0.47	0.47	1.121	1
	Trt:Gen	275.944	34.493	8	1.246	0.294	0.085	0.175	0.017	0.031	0.017	0.461	0.595
	Residuals	1301.15	27.684	47									
Man	Treatment	0.15	0.075	2	0.201	0.819	0.007	0.009	-0.027	-0.027	0.028	0.093	0.081
	Genotype	1.714	0.428	4	1.147	0.347	0.08	0.091	0.01	0.01	0.01	0.316	0.366

Mns	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's 'f	power
	Trt:Gen	2.365	0.296	8	0.792	0.613	0.11	0.121	-0.029	-0.028	-0.029	0.371	0.384
	Residuals	17.18	0.373	46									
Cel	Treatment	0.082	0.041	2	0.661	0.521	0.023	0.027	-0.011	-0.011	-0.012	0.168	0.162
	Genotype	0.285	0.071	4	1.145	0.347	0.079	0.089	0.01	0.009	0.01	0.312	0.365
	Trt:Gen	0.321	0.04	8	0.644	0.736	0.089	0.099	-0.048	-0.048	-0.049	0.331	0.312
	Residuals	2.925	0.062	47									

Mns	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's 'f	power
MnsT	Treatment	203.315	101.658	2	1.082	0.347	0.016	0.044	0.001	0.003	0.001	0.215	0.241
	Genotype	5740.827	1435.207	4	15.277	0	0.447	0.565	0.414	0.479	0.417	1.14	1
	Trt:Gen	2494.792	311.849	8	3.32	0.004	0.194	0.361	0.135	0.23	0.136	0.752	0.979
	Residuals	4415.316	93.943	47									
CrCel	Treatment	221.809	110.905	2	2.101	0.135	0.04	0.089	0.021	0.037	0.021	0.313	0.433
	Genotype	1794.133	448.533	4	8.498	0	0.325	0.441	0.284	0.341	0.287	0.889	0.999

Mns	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's 'f	power
	Trt:Gen	1231.225	153.903	8	2.916	0.011	0.223	0.352	0.145	0.209	0.147	0.737	0.958
	Residuals	2269.657	52.783	43									
Hem	Treatment	146.255	73.127	2	2.02	0.144	0.033	0.079	0.017	0.032	0.017	0.293	0.418
	Genotype	1810.358	452.589	4	12.5	0	0.412	0.515	0.376	0.426	0.379	1.031	1
	Trt:Gen	732.909	91.614	8	2.53	0.022	0.167	0.301	0.1	0.165	0.101	0.656	0.924
	Residuals	1701.672	36.206	47									

Mns	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's 'f	power
cel.hem	Treatment	0.113	0.056	2	1.693	0.196	0.047	0.075	0.019	0.024	0.019	0.284	0.357
	Genotype	0.465	0.116	4	3.481	0.015	0.194	0.249	0.137	0.148	0.139	0.576	0.864
	Trt:Gen	0.41	0.051	8	1.537	0.174	0.172	0.227	0.059	0.07	0.06	0.541	0.707
	Residuals	1.402	0.033	42									
ara.xyl	Treatment	0.02	0.01	2	2.212	0.121	0.052	0.086	0.028	0.038	0.029	0.307	0.453
	Genotype	0.1	0.025	4	5.569	0.001	0.264	0.322	0.214	0.228	0.217	0.688	0.979
	Trt:Gen	0.048	0.006	8	1.334	0.251	0.126	0.185	0.031	0.041	0.032	0.476	0.631

Mns	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
	Residuals	0.211	0.004	47									

Table G. 13 ANOVA test for the effect of treatment on the amount of matrix monosaccharides on the genotypes

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
Fuc	Bors	Treatment	0.013	0.006	2	0.32	0.734	0.066	0.066	-0.128	-0.128	-0.141	0.267	0.1
		Residuals	0.179	0.02	9									
	Gaheris	Treatment	0.046	0.023	2	0.463	0.644	0.093	0.093	-0.098	-0.098	-0.108	0.321	0.123

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
		Residuals	0.443	0.049	9									
	Galahad	Treatment	0.011	0.006	2	0.15	0.863	0.029	0.029	-0.15	-0.15	-0.165	0.173	0.073
		Residuals	0.37	0.037	10									
	Lamorak	Treatment	0.058	0.029	2	1.333	0.311	0.228	0.228	0.053	0.053	0.057	0.544	0.279
		Residuals	0.197	0.022	9									
	Tristan	Treatment	0.187	0.093	2	2.498	0.132	0.333	0.333	0.187	0.187	0.2	0.707	0.486

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
		Residuals	0.374	0.037	10									
Ara	Bors	Treatment	1.452	0.726	2	0.896	0.442	0.166	0.166	-0.018	-0.018	-0.019	0.446	0.199
		Residuals	7.297	0.811	9									
	Gaheris	Treatment	0.626	0.313	2	0.159	0.855	0.034	0.034	-0.163	-0.163	-0.18	0.188	0.074
		Residuals	17.672	1.964	9									
	Galahad	Treatment	1.052	0.526	2	0.492	0.625	0.09	0.09	-0.085	-0.085	-0.092	0.314	0.129

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
		Residuals	10.682	1.068	10									
	Lamorak	Treatment	4.647	2.323	2	3.9	0.06	0.464	0.464	0.326	0.326	0.345	0.931	0.679
		Residuals	5.361	0.596	9									
	Tristan	Treatment	0.781	0.39	2	0.414	0.672	0.076	0.076	-0.099	-0.099	-0.108	0.288	0.116
		Residuals	9.436	0.944	10									
Glc	Bors the Younger	Treatment	108.109	54.054	2	3.746	0.066	0.454	0.454	0.314	0.314	0.333	0.912	0.66

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
		Residuals	129.877	14.431	9									
	Gaheris	Treatment	8.296	4.148	2	0.269	0.77	0.056	0.056	-0.139	-0.139	-0.153	0.244	0.091
		Residuals	139.02	15.447	9									
	Galahad	Treatment	9.028	4.514	2	2.148	0.167	0.301	0.301	0.15	0.15	0.161	0.656	0.427
		Residuals	21.011	2.101	10									
	Lamorak	Treatment	80.877	40.439	2	3.198	0.089	0.415	0.415	0.268	0.268	0.286	0.843	0.588

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
		Residuals	113.808	12.645	9									
	Tristan	Treatment	67.923	33.961	2	12.368	0.002	0.712	0.712	0.636	0.636	0.655	1.573	0.994
		Residuals	27.46	2.746	10									
Xyl	Bors the Younger	Treatment	29.31	14.655	2	0.42	0.669	0.085	0.085	-0.107	-0.107	-0.118	0.306	0.116
		Residuals	313.747	34.861	9									
	Gaheris	Treatment	35.945	17.973	2	0.365	0.704	0.075	0.075	-0.118	-0.118	-0.131	0.285	0.107

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
		Residuals	443.232	49.248	9									
	Galahad	Treatment	141.225	70.612	2	4.947	0.032	0.497	0.497	0.378	0.378	0.397	0.995	0.789
		Residuals	142.743	14.274	10									
	Lamorak	Treatment	103.425	51.712	2	2.155	0.172	0.324	0.324	0.161	0.161	0.174	0.692	0.425
		Residuals	215.999	24	9									
	Tristan	Treatment	2.688	1.344	2	0.072	0.931	0.014	0.014	-0.166	-0.166	-0.183	0.12	0.061

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
		Residuals	185.433	18.543	10									
Man	Bors	Treatment	0.15	0.075	2	1.733	0.237	0.302	0.302	0.118	0.118	0.128	0.658	0.348
		Residuals	0.346	0.043	8									
	Gaheris	Treatment	3.621	1.81	2	1.114	0.37	0.198	0.198	0.019	0.019	0.02	0.497	0.239
		Residuals	14.633	1.626	9									
	Galahad	Treatment	0.46	0.23	2	2.076	0.176	0.293	0.293	0.142	0.142	0.152	0.644	0.415

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
		Residuals	1.107	0.111	10									
	Lamorak	Treatment	0.194	0.097	2	1.166	0.355	0.206	0.206	0.027	0.027	0.029	0.509	0.248
		Residuals	0.747	0.083	9									
	Tristan	Treatment	0.166	0.083	2	2.385	0.142	0.323	0.323	0.176	0.176	0.188	0.691	0.467
		Residuals	0.347	0.035	10									
Cel	Bors	Treatment	0.082	0.041	2	0.788	0.484	0.149	0.149	-0.037	-0.037	-0.04	0.418	0.18

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
		Residuals	0.47	0.052	9									
	Gaheris	Treatment	0.003	0.001	2	0.016	0.984	0.004	0.004	-0.196	-0.196	-0.218	0.06	0.052
		Residuals	0.75	0.083	9									
	Galahad	Treatment	0.073	0.036	2	0.515	0.613	0.093	0.093	-0.081	-0.081	-0.088	0.321	0.133
		Residuals	0.706	0.071	10									
	Lamorak	Treatment	0.052	0.026	2	1.322	0.314	0.227	0.227	0.051	0.051	0.055	0.542	0.277

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
		Residuals	0.176	0.02	9									
	Tristan	Treatment	0.29	0.145	2	1.758	0.222	0.26	0.26	0.104	0.104	0.112	0.593	0.358
		Residuals	0.823	0.082	10									
Mns T	Bors	Treatment	203.315	101.658	2	0.789	0.483	0.149	0.149	-0.036	-0.036	-0.04	0.419	0.18
		Residuals	1159.666	128.852	9									
	Gaheris	Treatment	1758.771	879.386	2	15.801	0.001	0.778	0.778	0.712	0.712	0.729	1.874	0.999

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
		Residuals	500.876	55.653	9									
	Galahad	Treatment	766.158	383.079	2	5.73	0.022	0.534	0.534	0.421	0.421	0.441	1.07	0.847
		Residuals	668.582	66.858	10									
	Lamorak	Treatment	374.432	187.216	2	3.335	0.082	0.426	0.426	0.28	0.28	0.298	0.861	0.607
		Residuals	505.178	56.131	9									
	Tristan	Treatment	421.495	210.748	2	1.333	0.307	0.21	0.21	0.049	0.049	0.053	0.516	0.281

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
		Residuals	1581.015	158.101	10									
CrCel	Bors	Treatment	221.809	110.905	2	2.121	0.176	0.32	0.32	0.157	0.157	0.169	0.687	0.419
		Residuals	470.559	52.284	9									
	Gaheris	Treatment	650.333	325.166	2	6.193	0.02	0.579	0.579	0.464	0.464	0.486	1.173	0.87
		Residuals	472.549	52.505	9									
	Galahad	Treatment	358.329	179.165	2	1.834	0.21	0.268	0.268	0.114	0.114	0.122	0.606	0.372

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
		Residuals	976.766	97.677	10									
	Lamorak	Treatment	89.618	44.809	2	1.857	0.236	0.382	0.382	0.16	0.16	0.176	0.787	0.357
		Residuals	144.795	24.133	6									
	Tristan	Treatment	69.433	34.716	2	1.524	0.269	0.253	0.253	0.08	0.08	0.087	0.582	0.314
		Residuals	204.988	22.776	9									
Hem	Bors	Treatment	146.255	73.127	2	1.363	0.304	0.233	0.233	0.057	0.057	0.062	0.55	0.284

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
		Residuals	482.755	53.639	9									
	Gaheris	Treatment	413.113	206.556	2	5.085	0.033	0.531	0.531	0.405	0.405	0.426	1.063	0.796
		Residuals	365.607	40.623	9									
	Galahad	Treatment	200.776	100.388	2	4.16	0.048	0.454	0.454	0.327	0.327	0.345	0.912	0.713
		Residuals	241.346	24.135	10									
	Lamorak	Treatment	300.691	150.345	2	5.678	0.025	0.558	0.558	0.438	0.438	0.46	1.123	0.839

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
		Residuals	238.288	26.476	9									
	Tristan	Treatment	101.332	50.666	2	1.356	0.301	0.213	0.213	0.052	0.052	0.056	0.521	0.285
		Residuals	373.676	37.368	10									

Table G. 14 ANOVA test for the effect of genotype on the amount of matrix monosaccharides in the different treatments

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's f	power
Fuc	C	Genotype	0.35	0.088	4	3.445	0.029	0.434	0.434	0.298	0.298	0.308	0.875	0.854

		Residuals	0.458	0.025	18								
	N-P	Genotype	0.371	0.093	4	3.799	0.032	0.559	0.559	0.397	0.397	0.412	0.884
		Residuals	0.293	0.024	12								
	P-C	Genotype	0.452	0.113	4	2.365	0.094	0.357	0.357	0.199	0.199	0.206	0.681
		Residuals	0.813	0.048	17								
Ara	C	Genotype	38.698	9.675	4	7.069	0.001	0.611	0.611	0.513	0.513	0.525	0.994
		Residuals	24.635	1.369	18								

	N-	Genotype	31.23	7.808	4	7.533	0.003	0.715	0.715	0.606	0.606	0.62	1.585	0.995
		Residuals	12.438	1.036	12								
	P-	Genotype	44.893	11.223	4	14.264	0	0.77	0.77	0.707	0.707	0.716	1.832	1
		Residuals	13.376	0.787	17								
Glc	C	Genotype	67.7	16.925	4	2.106	0.122	0.319	0.319	0.161	0.161	0.167	0.684	0.625
		Residuals	144.638	8.035	18								
	N-	Genotype	52.832	13.208	4	0.633	0.648	0.174	0.174	-0.094	-0.094	-0.101	0.459	0.207

		Residuals	250.307	20.859	12								
	P-	Genotype	73.058	18.265	4	8.57	0.001	0.668	0.668	0.579	0.579	0.59	0.999
		Residuals	36.23	2.131	17								
Xyl	C	Genotype	1635.088	408.772	4	15.699	0	0.777	0.777	0.719	0.719	0.728	1.868	1
		Residuals	468.69	26.038	18								
	N-	Genotype	862.911	215.728	4	7.74	0.003	0.721	0.721	0.613	0.613	0.628	1.606	0.996
		Residuals	334.482	27.873	12								

	P-	Genotype	1663.801	415.95	4	14.2	0	0.77	0.77	0.706	0.706	0.715	1.828	1
		Residuals	497.983	29.293	17									
Man	C	Genotype	1.714	0.428	4	0.515	0.726	0.103	0.103	-0.092	-0.092	0.097	0.338	0.175
		Residuals	14.977	0.832	18									
	N-	Genotype	0.866	0.217	4	4.084	0.026	0.577	0.577	0.421	0.421	0.435	1.167	0.903
		Residuals	0.636	0.053	12									
	P-	Genotype	0.583	0.146	4	1.488	0.253	0.271	0.271	0.085	0.085	0.089	0.61	0.461

		Residuals	1.567	0.098	16								
Cel	C	Genotype	0.285	0.071	4	1.738	0.186	0.279	0.279	0.114	0.114	0.118	0.532
		Residuals	0.738	0.041	18								
	N-	Genotype	0.336	0.084	4	1.385	0.297	0.316	0.316	0.083	0.083	0.088	0.426
		Residuals	0.728	0.061	12								
	P-	Genotype	0.12	0.03	4	0.35	0.841	0.076	0.076	-0.134	-0.134	-0.141	0.131
		Residuals	1.459	0.086	17								

Mns T	C	Genotype	5740.827	1435.207	4	33.577	0	0.882	0.882	0.85	0.85	0.856	2.732	1
		Residuals	769.386	42.744	18								
	N-	Genotype	3625.517	906.379	4	8.334	0.002	0.735	0.735	0.633	0.633	0.647	1.667	0.998
		Residuals	1305.123	108.76	12								
	P-	Genotype	5226.551	1306.638	4	9.489	0	0.691	0.691	0.607	0.607	0.618	1.494	1
		Residuals	2340.808	137.695	17								
CrCel	C	Genotype	1794.133	448.533	4	19.331	0	0.811	0.811	0.761	0.761	0.769	2.073	1

		Residuals	417.656	23.203	18								
	N-C	Genotype	440.37	110.092	4	0.942	0.476	0.255	0.255	0.015	-0.015	0.016	0.585	0.295
		Residuals	1286.201	116.927	11									
	P-C	Genotype	706.123	176.531	4	4.368	0.017	0.555	0.555	0.415	0.415	0.428	1.117	0.927
		Residuals	565.801	40.414	14									
Hem	C	Genotype	1810.358	452.589	4	11.82	0	0.724	0.724	0.653	0.653	0.663	1.621	1
		Residuals	689.236	38.291	18									

N-	Genotype	1737.366	434.341	4	15.81	0	0.841	0.777	0.777	0.787	2.296	1	
	Residuals	329.673	27.473	12									
P-	Genotype	2972.001	743	4	18.5	0	0.813	0.761	0.761	0.769	2.086	1	
	Residuals	682.763	40.163	17									

Table G. 15

Mns	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power

Cel/Hem	Treatment	0.113	0.056	2	1.693	0.196	0.047	0.075	0.019	0.024	0.019	0.284	0.357
	Genotype	0.465	0.116	4	3.481	0.015	0.194	0.249	0.137	0.148	0.139	0.576	0.864
	Trt:Gen	0.41	0.051	8	1.537	0.174	0.172	0.227	0.059	0.07	0.06	0.541	0.707
	Residuals	1.402	0.033	42									
Ara/Xyl	Treatment	0.02	0.01	2	2.212	0.121	0.052	0.086	0.028	0.038	0.029	0.307	0.453
	Genotype	0.1	0.025	4	5.569	0.001	0.264	0.322	0.214	0.228	0.217	0.688	0.979
	Trt:Gen	0.048	0.006	8	1.334	0.251	0.126	0.185	0.031	0.041	0.032	0.476	0.631
	Residuals	0.211	0.004	47									

Table G. 16

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's 'f	power

Cell/Hem	Bors the Younger	Treatment	0.113	0.056	2	2.097	0.179	0.318	0.318	0.155	0.155	0.166	0.683	0.415
		Residuals	0.243	0.027	9									
	Gaheris	Treatment	0.124	0.062	2	2.019	0.195	0.335	0.335	0.156	0.156	0.169	0.71	0.398
		Residuals	0.246	0.031	8									
	Galahad	Treatment	0.008	0.004	2	0.132	0.878	0.026	0.026	-0.154	-0.154	-0.169	0.162	0.07
		Residuals	0.299	0.029	10									
	Lamorak	Treatment	0.165	0.083	2	29.36	0.001	0.907	0.907	0.863	0.863	0.876	3.128	1

		Residuals	0.017	0.003	6								
	Tristan	Treatment	0.032	0.016	2	0.237	0.794	0.05	0.05	-0.146	0.146	-0.161	0.229	0.086
		Residuals	0.606	0.067	9								
Ara/Xyl	Bors the Younger	Treatment	0.02	0.01	2	1.439	0.289	0.241	0.241	0.067	0.067	0.072	0.564	0.296
		Residuals	0.063	0.007	9								
	Gaheris	Treatment	0.003	0.002	2	0.131	0.879	0.028	0.028	-0.169	0.169	-0.188	0.171	0.07
		Residuals	0.106	0.012	9								

Galahad	Treatment	0.039	0.019	2	8.325	0.007	0.625	0.625	0.53	0.53	0.55	1.29	0.952
	Residuals	0.023	0.002	10									
Lamorak	Treatment	0	0	2	0.049	0.953	0.011	0.011	-0.188	-0.188	-0.209	0.104	0.057
	Residuals	0.01	0.001	9									
Tristan	Treatment	0.001	0.001	2	0.586	0.575	0.105	0.105	-0.068	-0.068	-0.074	0.342	0.145
	Residuals	0.01	0.001	10									

Table G. 17

Mns	Treatment	Source	SS	MS	df	F	p	η2	part η2	ω2	part ω2	ε2	Cohen's'f	power
Cel/Hem	C	Genotype	0.465	0.116	4	3.087	0.042	0.407	0.407	0.266	0.266	0.275	0.828	0.808
		Residuals	0.677	0.038	18									
	N-	Genotype	0.156	0.039	4	5.348	0.012	0.66	0.66	0.521	0.521	0.537	1.395	0.964
		Residuals	0.08	0.007	11									
	P-	Genotype	0.301	0.075	4	1.516	0.255	0.318	0.318	0.103	0.103	0.108	0.683	0.465
		Residuals	0.644	0.05	13									
Ara/Xyl	C	Genotype	0.1	0.025	4	3.868	0.019	0.462	0.462	0.333	0.333	0.343	0.927	0.895
		Residuals	0.116	0.006	18									
	N-	Genotype	0.025	0.006	4	0.978	0.455	0.246	0.246	-0.005	-0.005	-0.006	0.571	0.307
		Residuals	0.075	0.006	12									
	P-	Genotype	0.023	0.006	4	5.075	0.007	0.544	0.544	0.426	0.426	0.437	1.093	0.962
		Residuals	0.02	0.001	17									

G4 Pictures

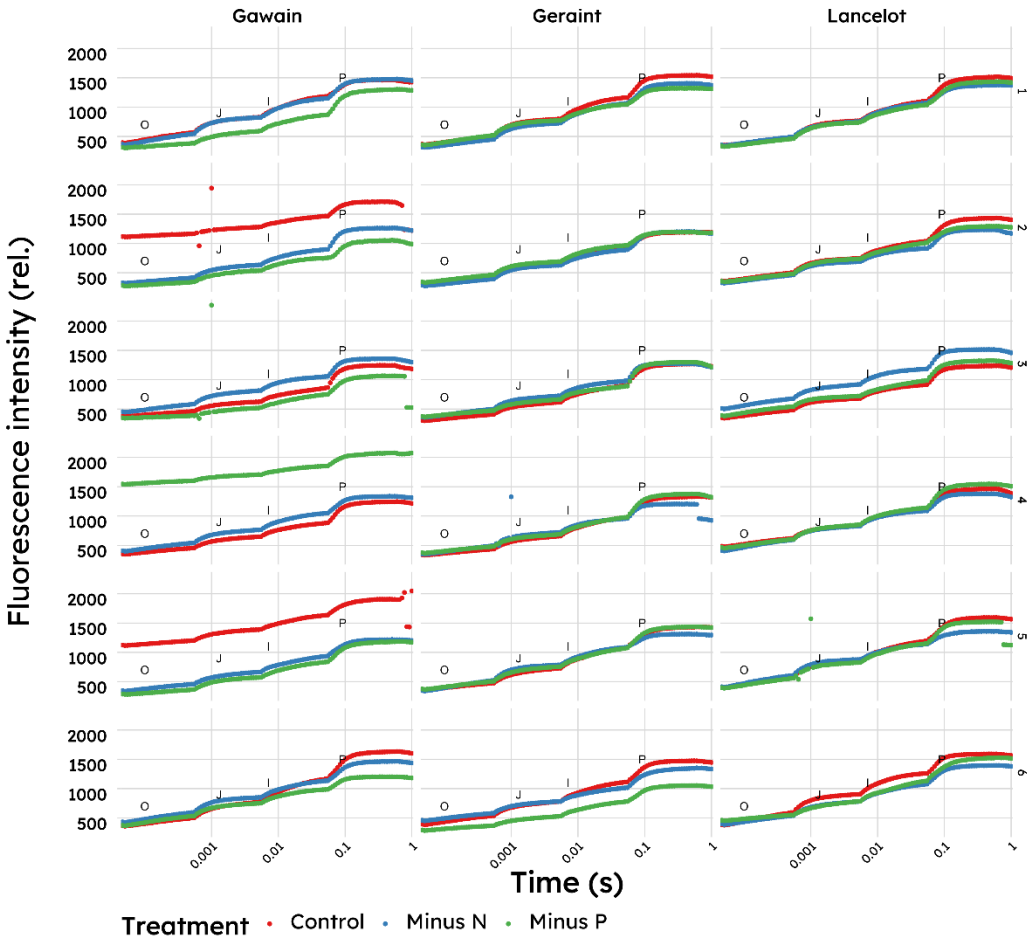


Figure G. 1 Chlorophyll a fluorescence induction curves in Gawain, Geraint and Lancelot

Chlorophyll a fluorescence induction curves were obtained after PSII inactivation (30 minutes of dark adaptation) of the youngest completely developed leaf. Measurement were performed at six time points (10, 20, 30, 40, 50 and 60 DOT respectively) during the duration of the treatment with limited nitrogen and limited phosphorus measured using a HandyPEA fluorimeter (Hansatech, UK).

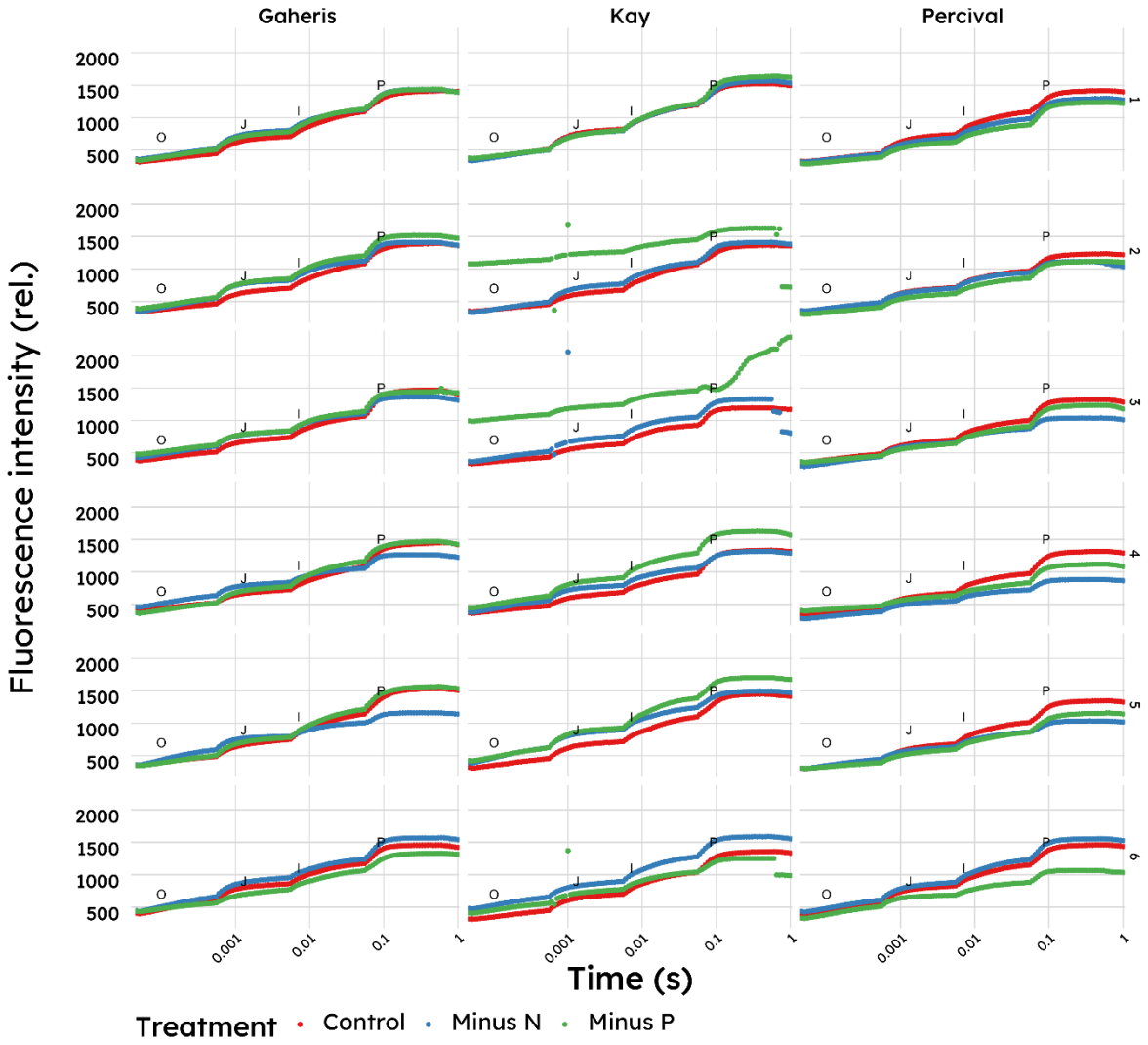


Figure G. 2 Chlorophyll a fluorescence induction curves in Gaheris, Kay and Percival

Chlorophyll a fluorescence induction curves were obtained after PSII inactivation (30 minutes of dark adaptation) of the youngest completely developed leaf. Measurement were performed at six time points (10, 20, 30, 40, 50 and 60 DOT respectively) during the duration of the treatment with limited nitrogen and limited phosphorus measured using a HandyPEA fluorimeter (Hansatech, UK).

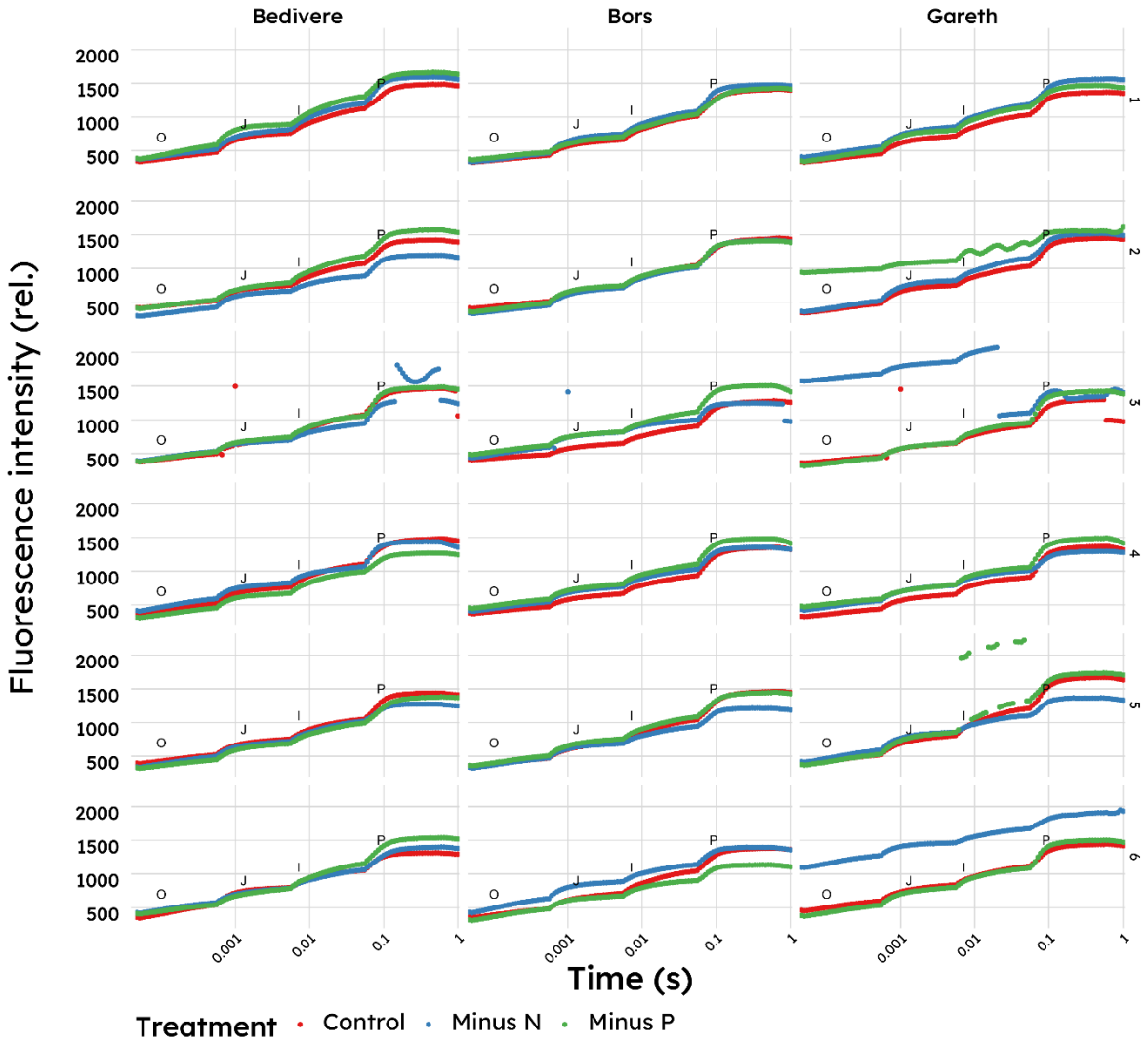


Figure G. 3 Chlorophyll a fluorescence induction curves in Bedivere, Bors and Gareth

Chlorophyll a fluorescence induction curves were obtained after PSII inactivation (30 minutes of dark adaptation) of the youngest completely developed leaf. Measurement were performed at six time points (10, 20, 30, 40, 50 and 60 DOT respectively) during the duration of the treatment with limited nitrogen and limited phosphorus measured using a HandyPEA fluorimeter (Hansatech, UK).

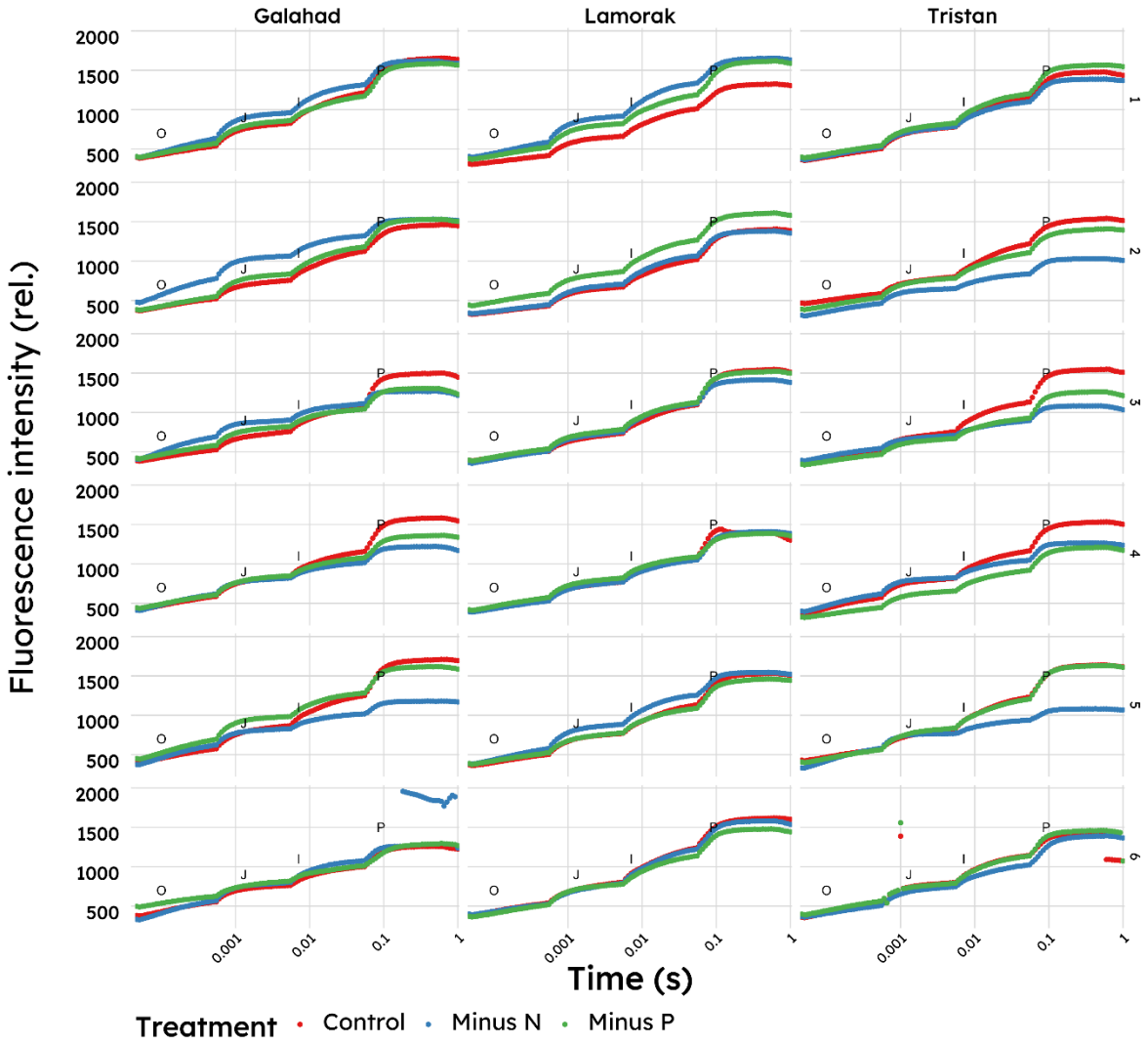


Figure G. 4 Chlorophyll a fluorescence induction curves in Galahad, Lamorak and Tristan

Chlorophyll a fluorescence induction curves were obtained after PSII inactivation (30 minutes of dark adaptation) of the youngest completely developed leaf. Measurement were performed at six time points (10, 20, 30, 40, 50 and 60 DOT respectively) during the duration of the treatment with limited nitrogen and limited phosphorus measured using a HandyPEA fluorimeter (Hansatech, UK).

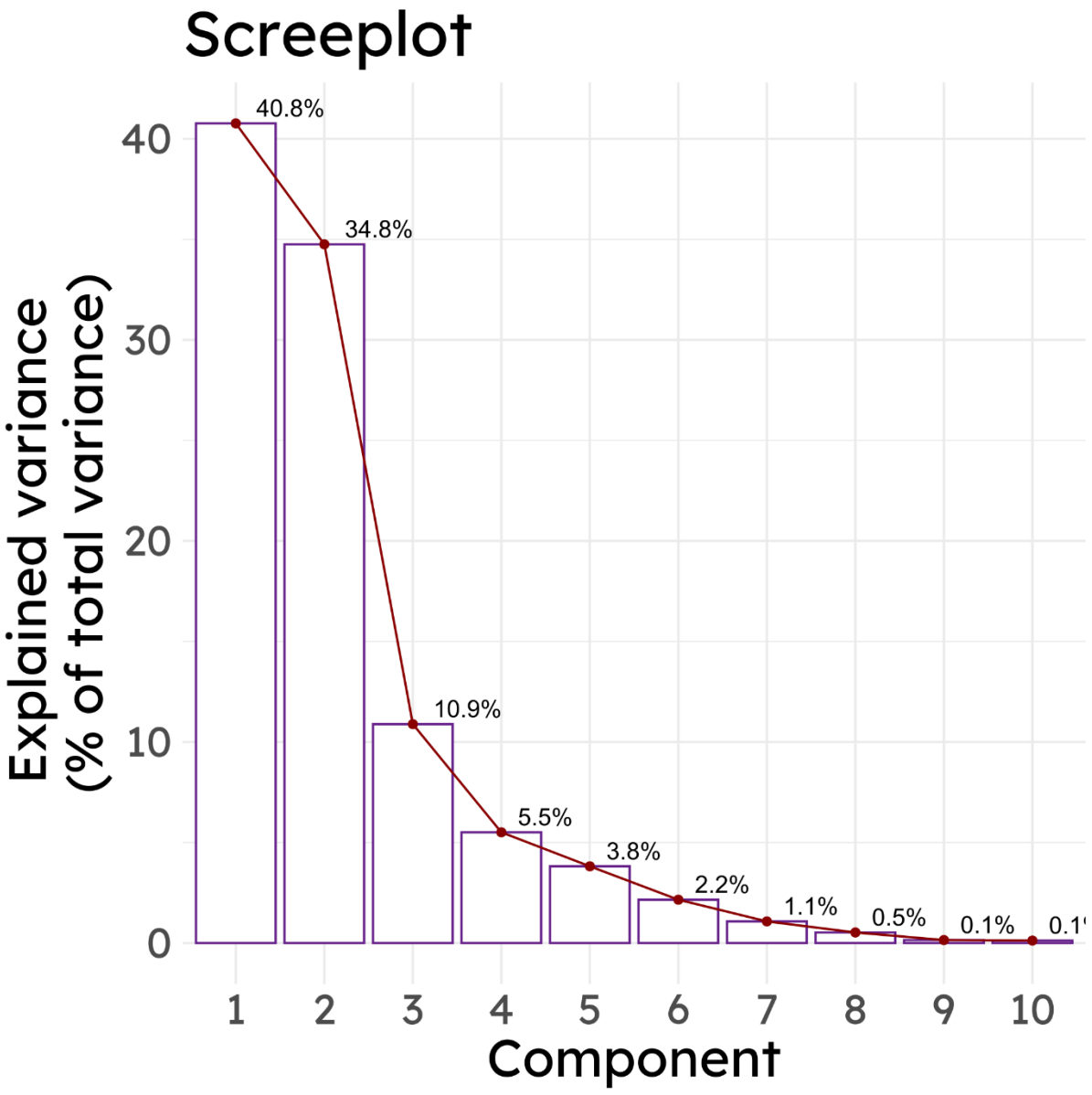


Figure G. 5 Screeplot of explained variance of the PCA components. A PCA was carried on the set of variates representing the chlorophyll a fluorescence parameters. The first two components explained 75.6% of the total variability.

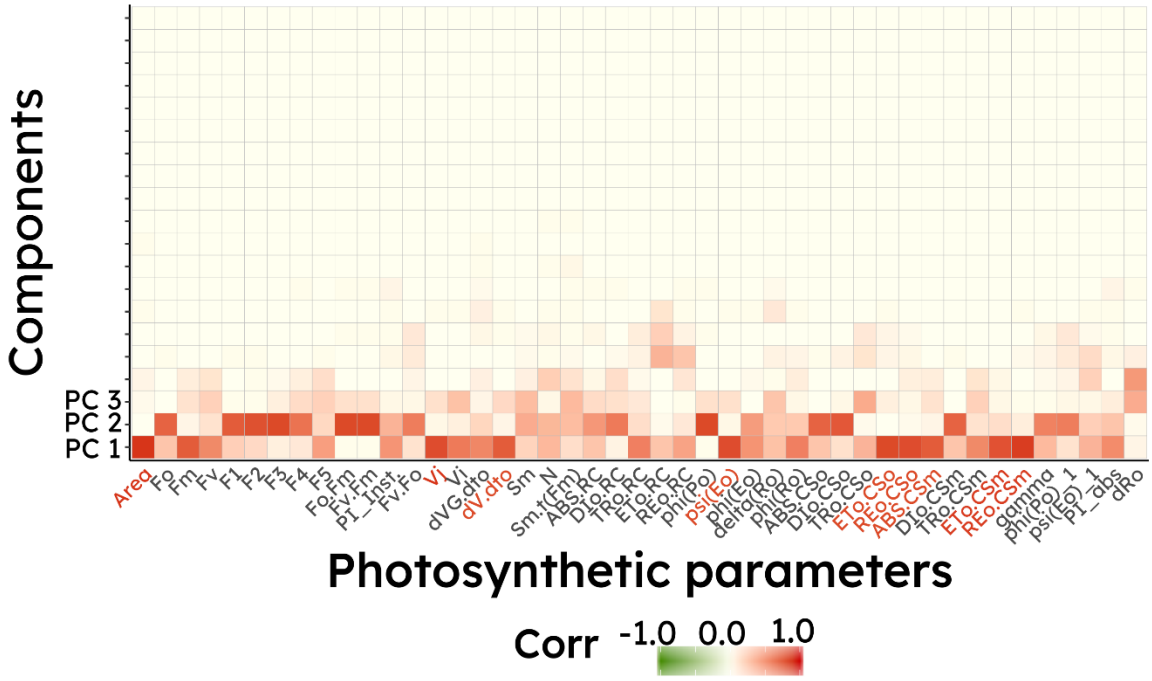


Figure G. 6 Quality of represention of the principal components.

The quality of representation of the variables on factor map is called cos2 (square cosine, squared coordinates) . The quality of the representation measures the degree of association between variable categories and a particular axis. The color scale represents the correlation. Variates that with the highest contribution to the first component are marked in red.

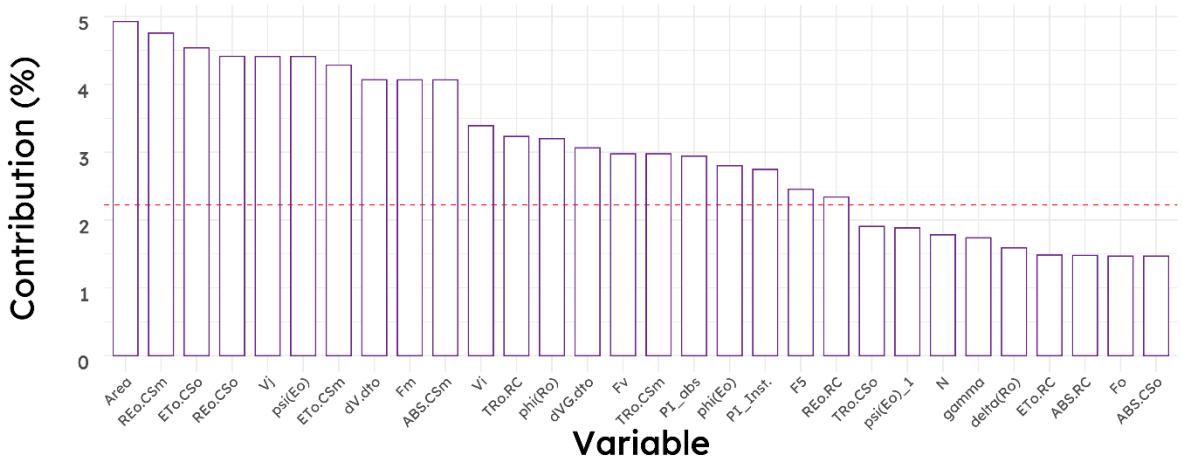


Figure G. 7 Contribution of the variables to the PC1.

The red dashed line on the graph above indicates the expected average contribution. If the contribution of the variables were uniform, the expected value would be 1/length(variables) = 1/10 = 10%. For a given component, a variable with a contribution larger than this cutoff could be considered as important in contributing to the component.

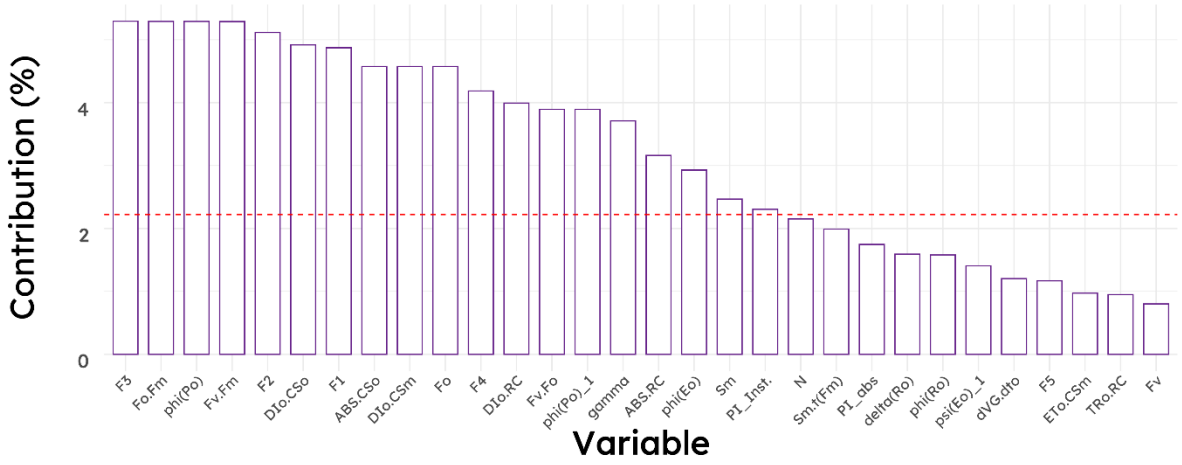


Figure G. 8 Contribution of the variables to the PC2.

The red dashed line on the graph above indicates the expected average contribution. If the contribution of the variables were uniform, the expected value would be

1/length(variables) = 1/10 = 10%. For a given component, a variable with a contribution larger than this cutoff could be considered as important in contributing to the component.

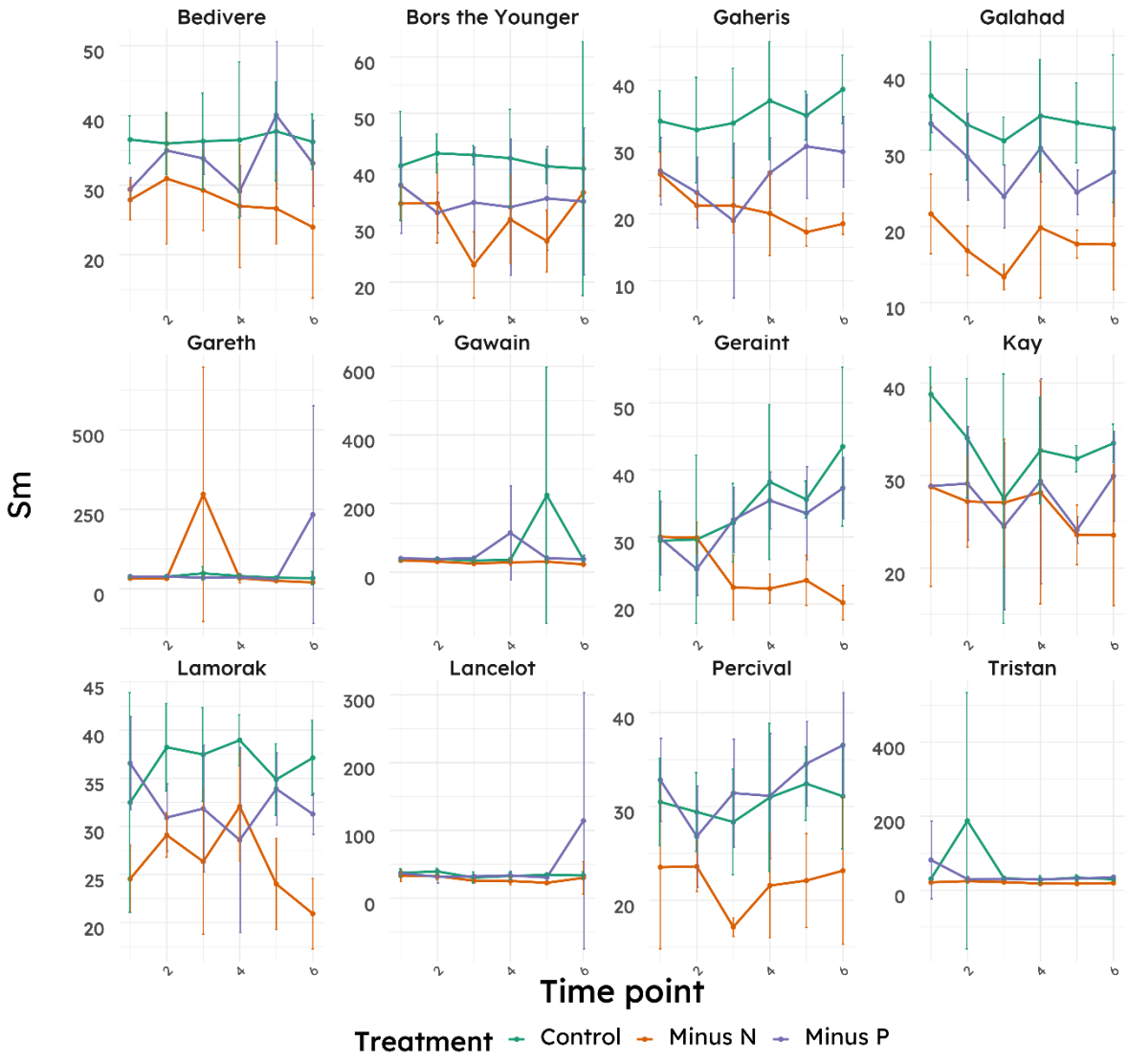


Figure G. 9 Normalised area under the fluorescence curve (Sm)

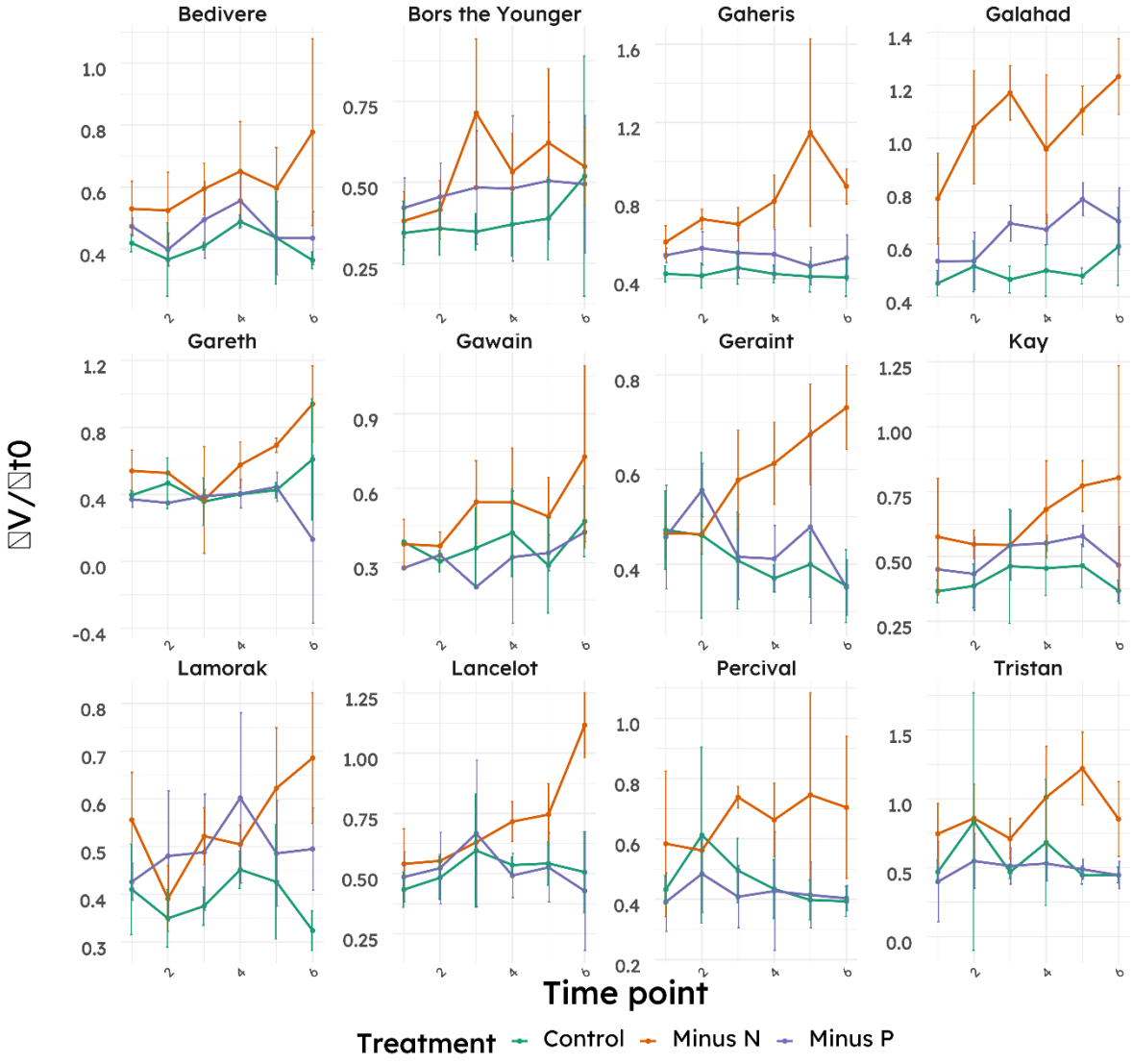


Figure G. 10 Slope of the fluorescence curve at time 0 (V0)

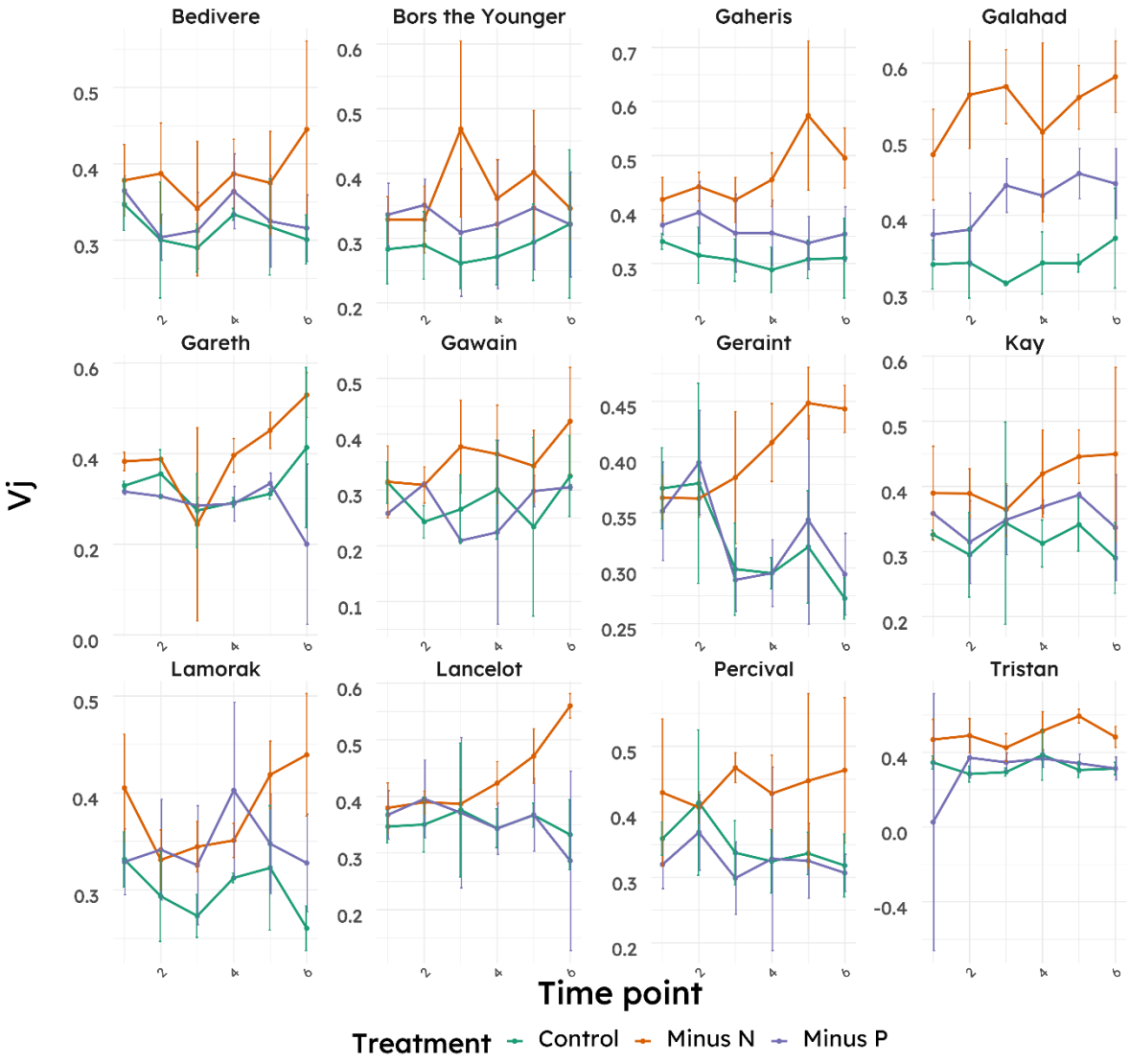


Figure G. 11 Relative fluorescence intensity at the J step (Vj)

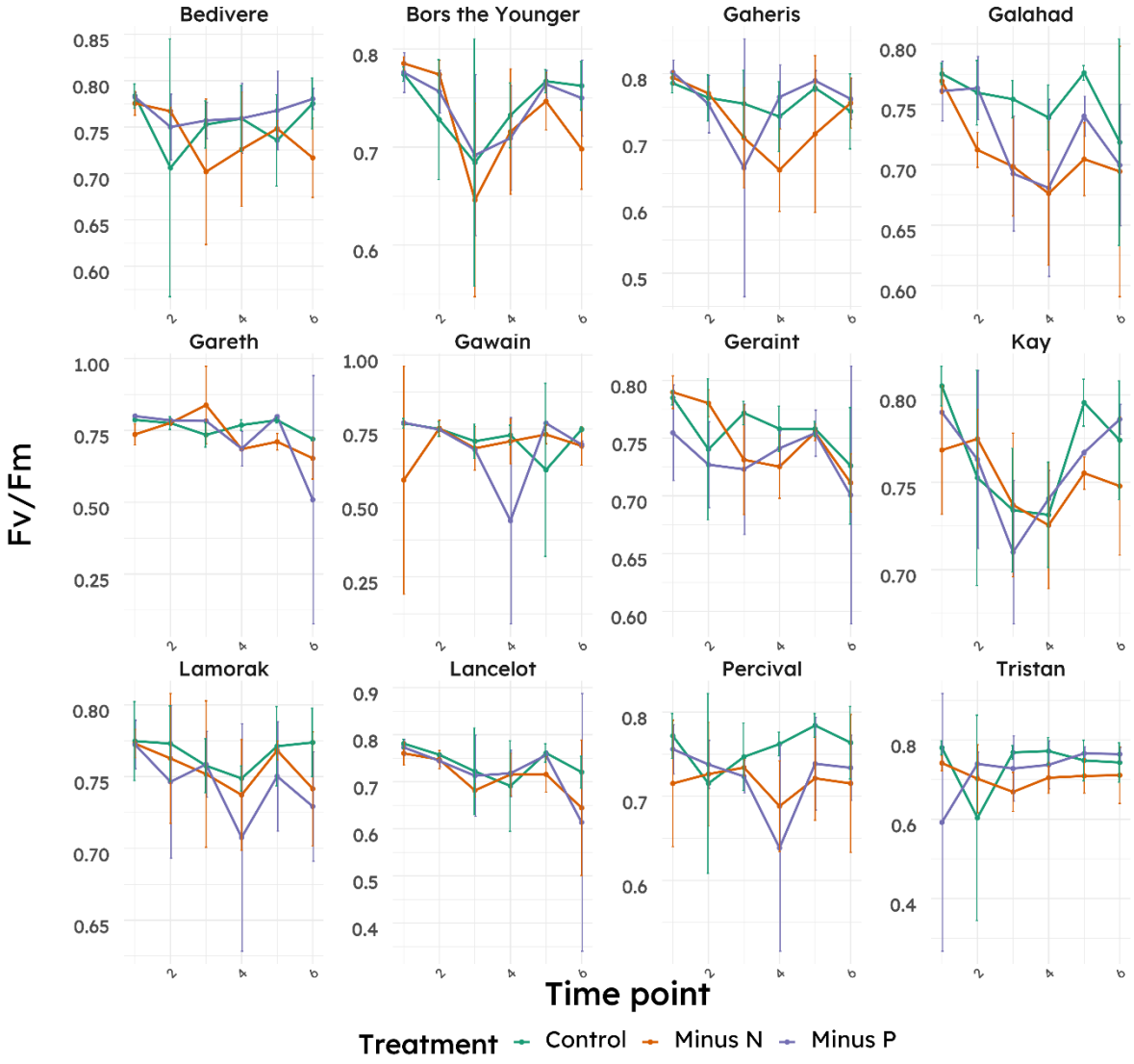


Figure G. 12 ratio between extremes of the fluorescence curve

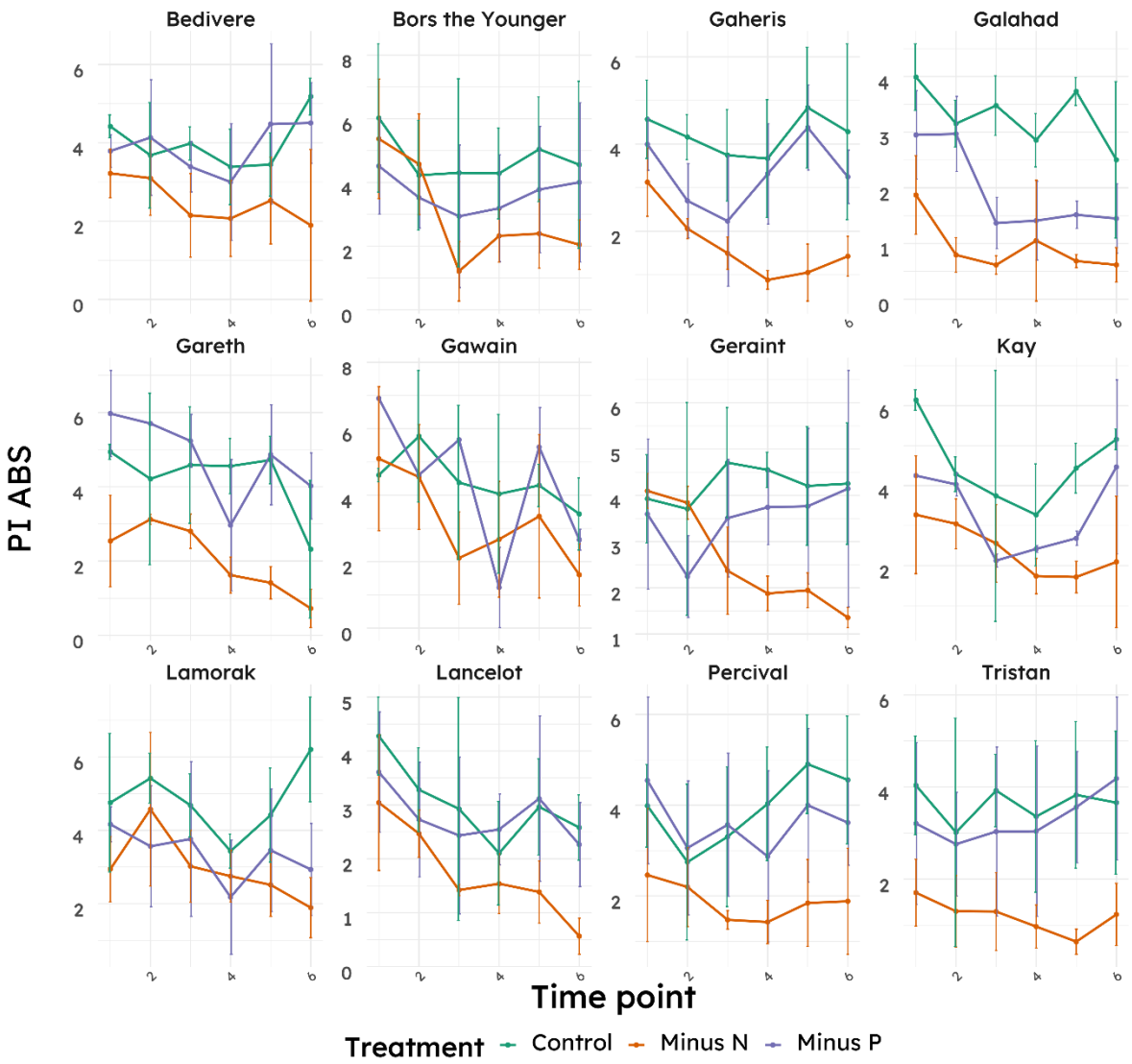


Figure G. 13 Performance index (PI)

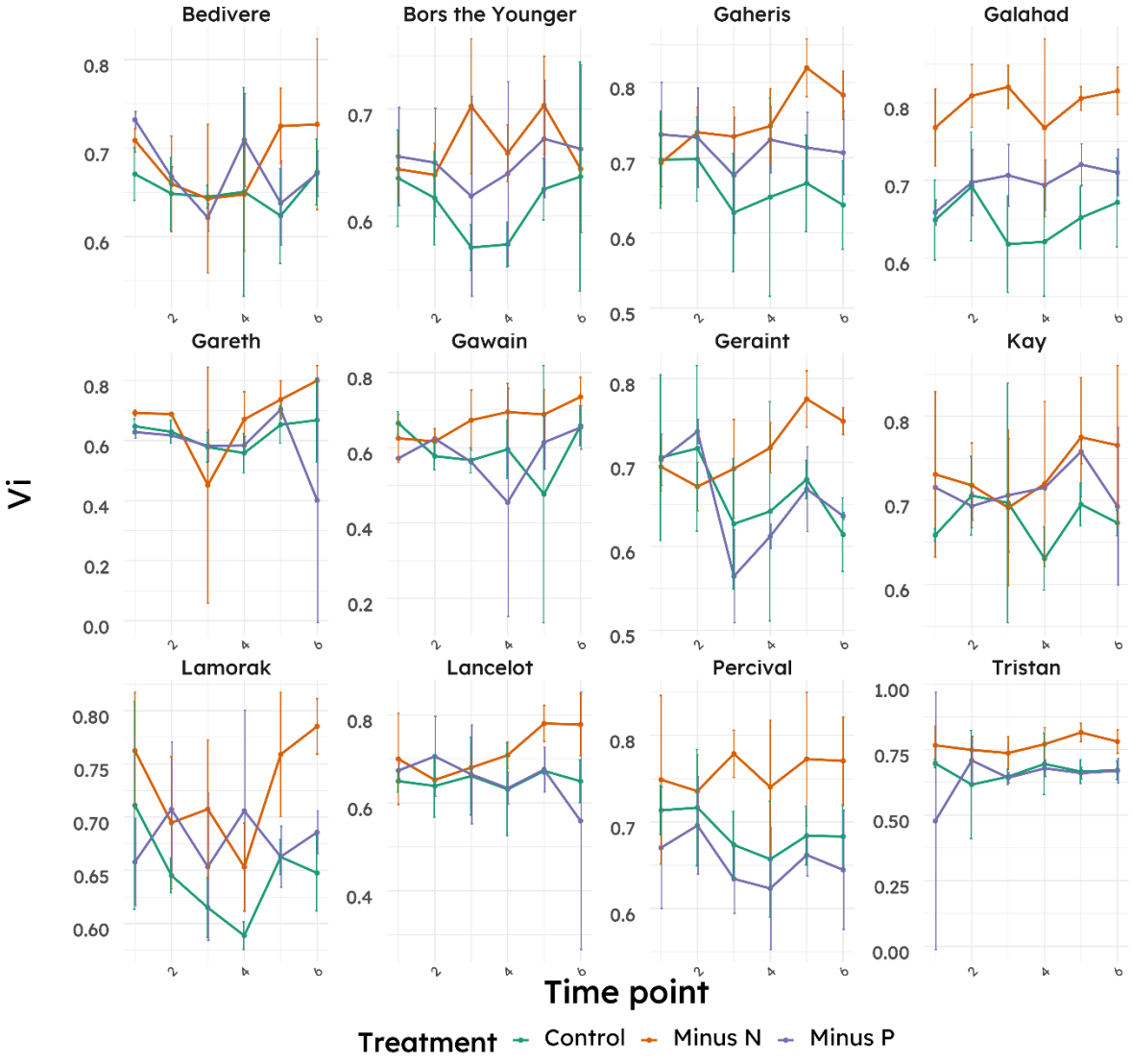


Figure G. 14 Realtive fluorescen intensity at the I step (Vi)

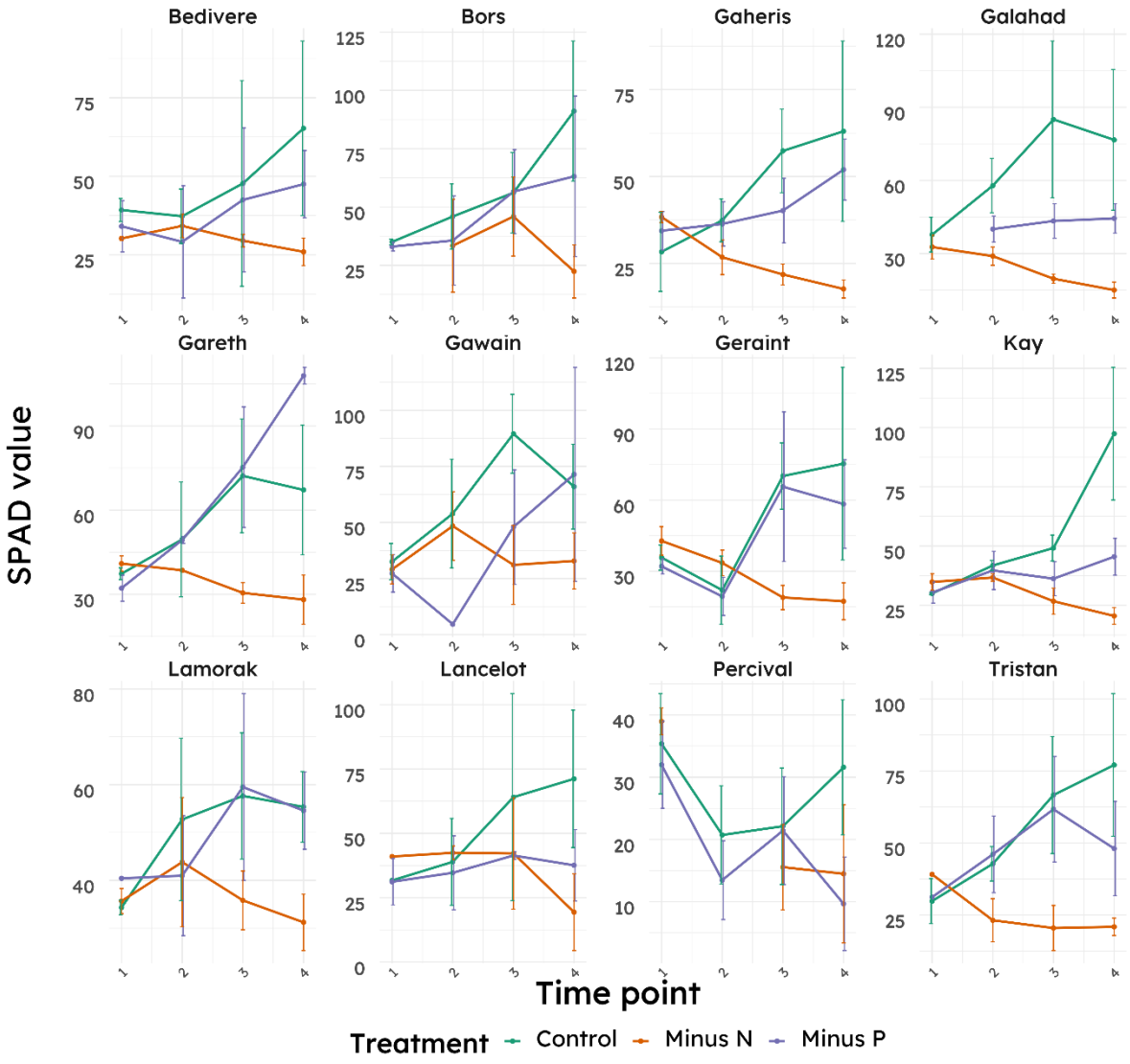


Figure G. 15 Chlorophyll content variation during the treatment.

Values on the x-axes are the chlorophyll content as determined by the SPAD meter. Points represent average between biological replicates of the measurement at 4 time points during the experiment. Bars represent the standard deviation of the mean. Different colors are used to represent the three experimental conditions.